UNIVERSITY OF CALIFORNIA

Santa Barbara

Mathematical Approaches to Understanding Mammalian Circadian Rhythms

A Dissertation submitted in partial satisfaction

of the requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

by

Peter Corbin St. John

Committee in charge:

Professor Francis J. Doyle III, Chair

Professor Samir Mitragotri

Professor Michelle A. O'Malley

Professor Linda R. Petzold

June 2015

The dissertation of Peter C. St. John is approved.

Professor Samir Mitragotri

Professor Michelle A. O'Malley

Professor Linda R. Petzold

Professor Francis J. Doyle III, Committee Chair

April 2015

Mathematical Approaches to Understanding Mammalian Circadian Rhythms

by

Peter C. St. John

Acknowledgement

First and foremost, I would like to think Frank Doyle III for taking the time to advise me and for providing numerous opportunities for me to present my research at conferences. I also thank the other members of my PhD thesis committee for their continued support and encouragement: Samir Mitragotri and Michelle O'Malley for keeping my work biologically relevant; and Linda Petzold for her expertise in numerical methods. I would also like to thank Linda for her excellent class in ODE integration, which greatly helped me understand and appreciate the algorithms I had largely taken for granted.

I would also like to thank the past and current members of the Doyle group, who have made my past five years so memorable. In no particular order, I thank Stephanie Taylor, for while we didn't actually overlap at UCSB, she nevertheless provided constant guidance and suggestions on the various manuscripts and presentations I prepared during my PhD thesis. I also thank Bharath Ananthasubramaniam and Panagiota (Pegy) Foteinou for their help getting me started when I first joined the groups and for many excellent lunchtime conversations. I also thank John Abel for his overall enthusiasm and specifically his help with stochastic methods, without which I may have taken a different path for my last two papers. I should not forget group members who have worked on other projects: Ravi Gondhalekar, Yongqiang Wang, and Felipe Núñez, who graciously put up with my constant questions about

spectral analysis and analytical methods.

I would also like to thank my fiancée Katherine LeVan - first in a professional context, for her depth of knowledge of classical statistics and assistance with many aspects of my research. Most importantly, of course, I thank her for her endless patience and support. Similarly, I thank my parents and brother, Christopher, for always being there when I need them.

Finally, I thank the various funding sources who have made this project possible. Most importantly, I appreciate the generous fellowship given to me by the Mitsubishi Chemical Corporation, which has given me the freedom to pursue whichever project caught my enthusiasm. I also thank financial support from the National Institutes of Health/National Institute of General Medical Sciences under award number 1R01GM096873-01 and from the Institute for Collaborative Biotechnologies through grant W911NF-09-0001 from the U.S. Army Research Office.

Vita of Peter C. St. John

Contact Information

Department of Chemical Engineering
University of California, Santa Barbara
Santa Barbara, CA 93106-5080

Phone: (508) 494-2474
E-mail: pstjohn@engineering.ucsb.edu
Office: Engineering II, Rm. 1508

Education

University of California, Santa Barbara — **September 2010 - June 2015**
Ph.D. Candidate, Department of Chemical Engineering — Santa Barbara, California

Tufts University — **September 2006 - May 2010**
BS, Chemical and Biological Engineering — Medford, Massachusetts

Honors and Awards

CAST Student Travel Grant	**September 2014**
Society for Research on Biological Rhythms Research Merit Award	**June 2014**
Best Poster, Center for Chronobiology Symposium, UCSD	**February 2014**
1st Place, SRBR Logo Competition	**January 2014**
Mitsubishi Chemical Fellowship Recipient	**2012-2015**
UCSB Scienceline 2011-2012 Life Science Outstanding Answerer	**June 2012**
National Science Foundation GRFP Honorable Mention	**April 2011**
Class of 1947 Victor Prather Prize	**May 2010**
Max Tischler Prize Scholarship	**May 2009**
Elected to Tau Beta Pi	**September 2008**

Publications

St. John, P.C. and F.J. Doyle III. Quantifying stochastic noise in cultured circadian reporter cells (2015), In review.

Abel, J.H., Widmer, L.A., **St. John, P.C.**, Stelling, J., and F.J. Doyle III. A Coupled Stochastic Model Explains Differences in Circadian Behavior of Cry1 and Cry2 Knockouts (2015) *IEEE Life Sciences Letters*, in press.

St. John, P.C., Taylor, S.R., Abel, J.H., and F.J. Doyle III. Amplitude metrics for cellular circadian bioluminescence reporters (2014) *Biophysical Journal*, 107 (11) pp. 2712-2722

St. John, P.C., Hirota, T., Kay, S.A., and F.J. Doyle III. Spatiotemporal separation of PER and CRY posttranslational regulation in the mammalian circadian clock (2014) *PNAS*, 111 (5) pp. 2040-2045.

Yang, R., Rodriguez-Fernandez, M., **St. John, P.C.,** and F.J. Doyle III. Chapter 8 – Systems Biology (2014) In E. Carson and C. Cobelli (Eds.) *Modelling Methodology for Physiology and Medicine, 2nd Edition*, pp. 159-187.

St. John, P.C., and F.J. Doyle III. Estimating confidence intervals in predicted responses for oscillatory biological models (2013) *BMC Systems Biology* 7:71.

Hirota, T., Lee, J.W., **St. John, P.C.**, Sawa, M., Iwaisako, K., Noguchi, T., Pongsawakul, P.Y., Sonntag, T., Welsh, D.K., Brenner, D.A., Doyle, F.J. III, Schultz, P.G., Kay, S.A., Identification of small molecule activators of cryptochrome (2012) *Science*, 337 (6098) pp. 1094-1097.

Murphy, A.R., **St. John, P.C.**, Kaplan, D.L. Modification of silk fibroin using diazonium coupling chemistry and the effects on hMSC proliferation and differentiation (2008) *Biomaterials*, 29 (19), pp. 2829-2838.

Contributed Talks

St. John, P.C., and F.J. Doyle III. November 2014. Development of Amplitude Response Curves for Single-Cell and Population-Level Circadian Systems. Presented at the 2014 AIChE Annual Meeting, Atlanta, GA

St. John, P.C., and F.J. Doyle III. June 2014. Amplitude metrics for uncoupled cellular circadian bioluminescence reporters. Presented at the Society for Research on Biological Rhythms Meeting, Big Sky, MT.

St. John, P.C., and F.J. Doyle III. October 2012. Cryptochrome balancing for period control: mathematical insights into circadian clock design. Presented at the Model-based Analysis and Control of Cellular Processes Workshop, Purdue University, West Lafayette, IN.

Poster Presentations

St. John, P.C., and F.J. Doyle III. November 2014. Development of Amplitude Response Curves for Single-Cell and Population-Level Circadian Systems. Presented at the 2014 AIChE Annual Meeting, Atlanta, GA

St. John, P.C., and F.J. Doyle III. February 2014. Spatiotemporal separation of PER and CRY post-translation regulation. Presented at the UCSD Center for Chronobiology Symposium, La Jolla, CA.

St. John, P.C., T. Hirota, S.A. Kay, and F.J. Doyle III. July 2013. Estimating confidence intervals in model predictions to determine plausible mechanisms for small molecule modifiers. Presented at the Chronobiology Gordon Research Conference, Newport, RI.

St. John, P.C., and F.J. Doyle III. February 2013. Predictive confidence intervals from mathematical circadian models. Presented at the UCSD Center for Chronobiology Symposium, La Jolla, CA.

St. John, P.C., T. Hirota, S. Kay, and F.J. Doyle III. May 2012. Cryptochrome balancing for period control. Presented at the Meeting of the Society for Research on Biological Rhythms, Destin, FL.

St. John, P.C., and F.J. Doyle III. February 2012. Perturbation analysis of circadian clock degradation. Presented at the UCSD Center for Chronobiology Symposium, La Jolla, CA.

St. John, P.C., M. Rodriguez-Fernandez, and F.J. Doyle III. June 2011. Advanced global optimization and sensitivity techniques for analyzing deterministic circadian models. Presented at the 12th Annual UC Systemwide Bioengineering Symposium, Santa Barbara, CA.

Seminars

St. John, P.C. May 2013. Sensitivity analysis in the study of mammalian circadian

rhythms. Presented at the UCSB NSF-IGERT systems biology seminar series, Santa Barbara, CA.

Research Experience

University of California, Santa Barbara **September 2010 - June 2015**
Ph.D. Candidate Santa Barbara, California
Computational analysis of the mammalian circadian clock, with a focus on elucidating the functional design consequences of the underlying genetic regulatory network.
Advisor: Francis J. Doyle III
Department: Chemical Engineering

Tufts University **September 2009 - May 2010**
Senior Honors Thesis Medford, Massachusetts
Catalytic Hydrodechlorination of 2-Chlorophenol using Viral-templated Palladium Nanoparticles
Advisor: Hyunmin Yi
Department: Chemical and Biological Engineering

University of California, Los Angeles **June 2009 - August 2009**
UCLA NanoCER Program (NSF REU) Los Angeles, California
Encapsulation Efficiencies and Release Rates from Water-in-Oil-in-Water Nanoemulsions
Advisor: Timothy Deming
Department: Bioengineering

Tufts University **June 2008 - August 2008**
Tufts Summer Scholars Medford, Massachusetts
Hydrodechlorination of 2-Chlorophenol with Palladium Nanoparticles on Genetically Modified Tobacco Mosaic Virus Scaffolds
Advisor: Hyunmin Yi
Department: Chemical and Biological Engineering

Tufts University **October 2007 - May 2008**
Undergraduate Research Credit Medford, Massachusetts
Chemically Modified Silk Fibroin Based Scaffolds for Bone Tissue Engineering
Advisor: David Kaplan
Department: Biomedical Engineering

Teaching Experience

University of California, Santa Barbara **January 2013 - December 2013**
Teaching Assistant, ChE132c Santa Barbara, California
Helped teach undergraduate statistics for two subsequent years: gave three lectures, held office hours and review sessions, and graded homeworks.

University of California, Santa Barbara **March 2012 - June 2012**
Teaching Assistant, ChE180a Santa Barbara, California
Designed and ran experiments for the junior laboratory course. Also helped in grading student reports.

Community Involvement

Peer Review **January 2014 - present**
Reviewer for Biophysical Journal; IEEE Control Systems Society Conference; 21st International Symposium on Mathematical Theory of Networks and Systems

Scienceline "Ask A Scientist" **December 2010 - April 2015**
Answers science and engineering questions posed by students and teachers from local K-12 schools.
Website: www.scienceline.ucsb.edu

UCSB Discover Engineering Weekend **May 2011**
Helped organize and run a weekend for local high school students to learn basic engineering principles and apply their knowledge to build miniature alternative energy cars.

Mentoring Experience

Amanda Luan, Undergraduate Student, ICB SSB URAP **June 2014 - April 2015**
Lukas Widmer, Masters Student, ETH Zurich **April 2012 - February 2013**
Andrew Barisser, Rotation Student, BMSE UCSB **September 2012 - December 2012**

Abstract

Mathematical Approaches to Understanding

Mammalian Circadian Rhythms

by

Peter Corbin St. John

Nearly all life on earth exists in a periodic environment, in which important factors like sunlight and temperature change predictably with a twenty-four hour cycle. As a process which only reacts to the current state of a periodic signal will constantly suffer a phase lag, organisms have developed a natural feedforward controller to predict upcoming environmental changes. Such a system allows an organism to align their behavior to the correct phase of the day/night cycle and ease transitions between times of energy abundance and energy scarcity. These daily changes in physiology are known as circadian rhythms and are coordinated by intricate genetic regulatory networks.

Over evolutionary timescales, nearly all aspects of gene expression have been coupled to the day/night cycle. As a result, circadian rhythms are essential to maintaining metabolic homeostasis, DNA repair, cell cycling, and other important cellular processes. Since modern societies have deviated from their evolutionary prescribed sleep and feeding schedules, disturbances to circadian gene expression have grown more common. Beyond acute effects on performance and fatigue, compromised cir-

cadian rhythms have been linked to chronic issues such as the onset of metabolic disease or increased cancer risk. Since circadian rhythms can be damped by factors such as jet lag, shift work, and high fat diets, there has been recent interest in developing pharmacological or behavioral therapies which might restore normal circadian rhythms.

This thesis uses techniques from dynamic systems to model circadian oscillations at different scales. First, a mathematical model of the core circadian feedback loop is developed in order to explain a novel small molecule modulator, KL001. Through this mathematical model, we gain new insight into how the two isoforms of cryptochrome (CRY1 and CRY2) interact to control the period. The identifiability of parameters and parametric sensitivities in oscillatory models is investigated next, and a dynamic optimization technique using collocation methods and nonlinear programming is shown to be able to efficiently bootstrap confidence intervals in such parameters. This technique is then applied to a set of three circadian models in order to identify mechanisms which are able to differentiate between the effect of two small molecule regulators, even across differences in parameter values and kinetic assumptions. Next, the effect of finite-duration perturbations on clock amplitude and synchrony is explored. New techniques and sensitivity analyses are developed which allow the effect of transient perturbations on the clock to be efficiently calculated without the need for computationally intensive stochastic simulations. Finally, the effect of clock perturbations on stochastic noise is investigated by fitting the damping rate of cultured cellular reporters. Using genome-wide siRNA knockdown screens, we are able to gain fundamental insight into design principles of circadian oscillations.

Contents

List of Figures

List of Models

List of Tables

Chapter 1

Introduction

Since life's first beginnings, nature has been seeking optimal solutions to surviving in a constantly changing environment. One key adaptation towards this goal, as evidenced by the ubiquity with which it is found throughout the tree of life, is the ability of an organism to anticipate daily changes to its environment. Such rhythms are known as *circadian*, and are classified by a number of defining characteristics [5]:

- They are *endogenous*, meaning they continue to oscillate even when the organism is isolated from its environment.
- They are temperature compensated, maintaining a consistent period for moderate changes in average temperature.
- They are *entrainable*, meaning they can adjust their phase and period in response to a changing in environmental signal.
- They oscillate with a roughly 24 hour period.

In this thesis, we apply techniques from systems biology in order to gain a more thorough understanding of how circadian rhythms control mammalian physiology,

and how these rhythms might be altered through pharmacological therapies. Systems biology is a multidisciplinary field, drawing from biology, mathematics, and computer science. In this chapter, we first provide a background on the biology of circadian rhythms, followed by an introduction to the mathematical techniques which have been used to study them. Some computational considerations in the estimation of model parameters are also discussed.

1.1 Biological background

1.1.1 Dynamics of gene expression

In the central dogma of molecular biology, genetic information is passed from DNA to mRNA to protein. While each cell in a multi-celled organism contains nearly identical DNA, the rate with which different genes are expressed results in substantial differences in cell type and function. These differences are the result of complex gene transcription networks, in which the production of some proteins - known as transcription factors - activate or repress the production of other genes. Similar to electrical circuits, network motifs such as positive and negative feedback arise from the interconnection of transcription factors.

Gene regulatory networks are essential to biological function. Cells maintain energy homeostasis by carefully balancing the flux of metabolites into energy storage and energy production pathways. These reactions are catalyzed by numerous enzymes, the activities of which must fluctuate to ensure the proper allocation of resources within a cell. Enzyme production, degradation, and activation are controlled by transcription factors, and thus environmental conditions can be processed by complex network interactions to regulate metabolic function. Circadian rhythms play a

major role in the regulation of metabolic cycles, as periodic shifts in energy availability require a similar periodic shifts in gene transcription.

1.1.2 Evolutionary history

By looking at the evolutionary history of circadian rhythms, Edgar and colleges [6] demonstrated that cellular timekeeping may have began as a response to the Great Oxidation Event occurring $\approx$ 2.5 billion years ago. Due to increased oxygen levels, the generation of reactive oxygen species in the cell can lead to oxidation of key biomolecules, prompting cells to express scavenging mechanisms during times of high UV light exposure. Over time, these pathways coupled to transcription-translation feedback loops in individual organisms, giving rise to a diverse set of circadian regulatory pathways (figure 1.1, from [6]).

1.1.3 Circadian rhythms in mammals

Tissue-level clocks

In mammals, circadian rhythms are organized in a hierarchical structure, in which inputs from the environment are processed separately by different tissue-level clocks (figure 1.2). Light, the primary entraining factor, is directed towards the suprachiasmatic nucleus (SCN), a tissue in the brain that serves as the body's master pacemaker [7]. In the SCN, approximately $20,000$ cells communicate via intercellular coupling to determine a precise internal time, which coordinates rhythms in locomotor activity [8].

Peripheral tissues are responsible for responding to changes daily changes in food intake [9]. Liver and adipose tissue, for instance, maintain robust circadian oscillations [10], which respond mainly to temperature and food resetting cues [11].

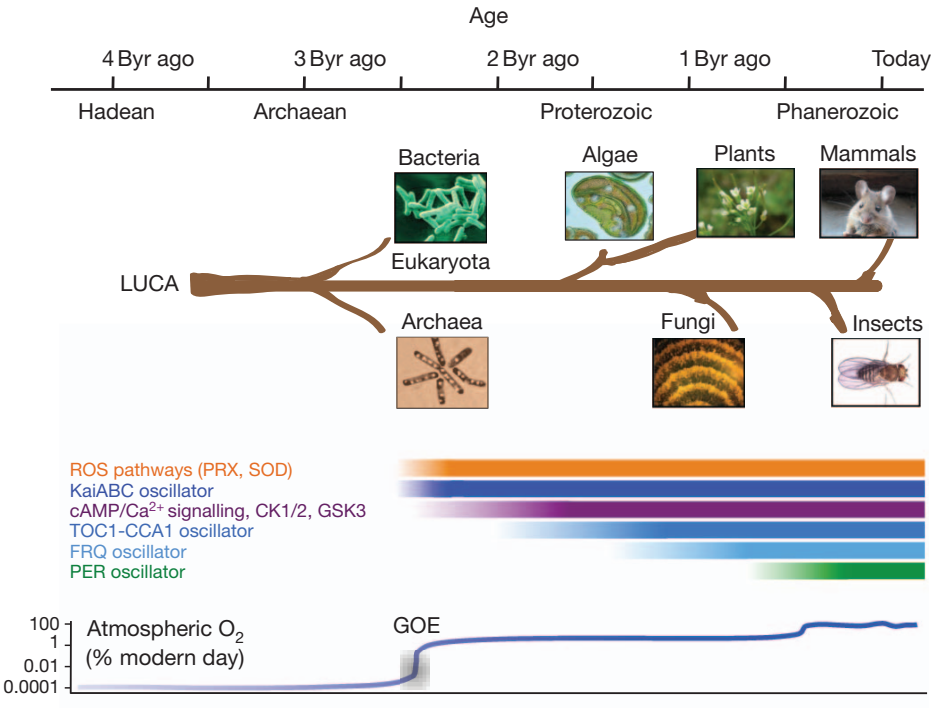


Figure 1.1: Reactive oxygen species are conserved markers of circadian rhythms. A timeline of the appearance of the main circadian feedback loops demonstrates that the rhythmic protection against light-mediated oxidation is a conserved rhythm across many evolutionary trees. Figure reprinted from [6].

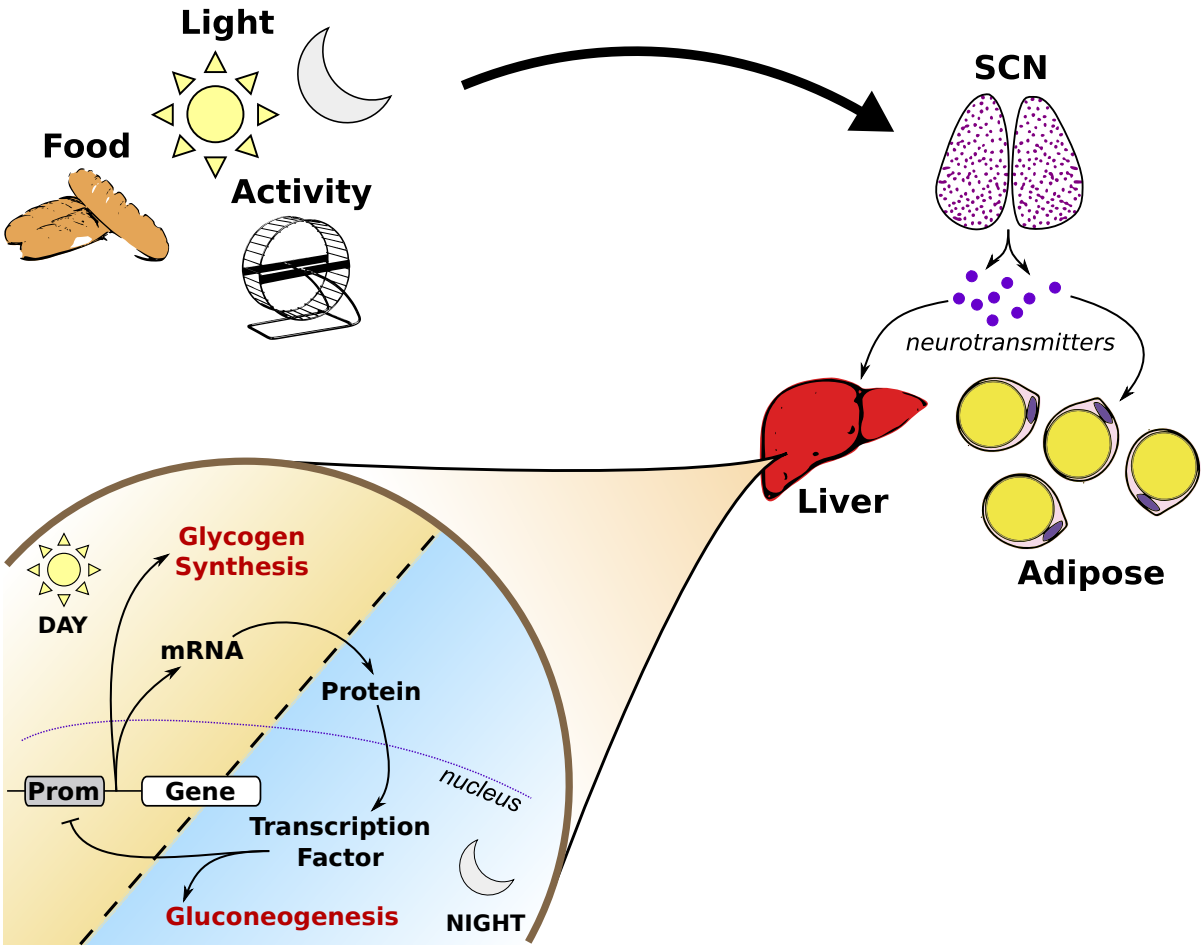


Figure 1.2: A biological feed forward controller. Circadian rhythms are organized in a hierarchical fashion, in which inputs from the environment are processed by different tissue-level clocks. At the single-cell level, rhythms are generated via a time-delayed transcription-translation negative feedback loop. Environmental inputs are used to predict upcoming environmental conditions by speeding or slowing the oscillatory cycle, matching the body's internal phase with the external environment.

Such clocks are important to maintaining metabolic health, as knockout experiments demonstrate that compromised circadian in peripheral tissues leads to many disorders, including diabetes and obesity [12, 13]. Additionally, feeding cycles have been shown to have a profound impact on circadian oscillations in peripheral tissues. High-fat or out of phase feeding leads to a reprogramming of oscillatory genes and metabolites [14, 15], often leading to metabolic disease.

Circadian oscillations at the single-cell level

Circadian rhythms are generated at the single-cell level through genetic regulatory networks with inherent time-delayed negative feedback. Transcription factors CLOCK and BMAL1, peaking in the early night, activate expression of EBOX genes period (*Per1* and *Per2*) and cryptochrome (*Cry1* and *Cry2*)[1]. *Per* and *Cry* protein products, PER and CRY, peaking during the early day, form a heterodimeric complex to cross the nuclear membrane [16]. However, since PER is present in the cytoplasm in much lower quantities than CRY, it is stochiometrically-limiting for nuclear entry [17]. Once inside the nucleus, the PER-CRY complex dissociates, and CRY represses CLOCK-BMAL1 mediated activation of EBOX genes, leading to rhythmic gene expression. Several time delays contribute to the 24 hour oscillatory period, highlighted in figure 1.3.

Additional components, namely the ROR and REV-ERB families of genes, add a layer of positive feedback to the clock regulatory loop. This addition results in more reliable clock oscillations, as the bistability induced by the positive feedback stabilizes the negative-feedback induced oscillations [18]. A schematic of the core feedback in mammalian circadian rhythms is shown in figure 1.4.

[1]In this thesis, we denote mRNA species using italics, and protein products using all capital letters.

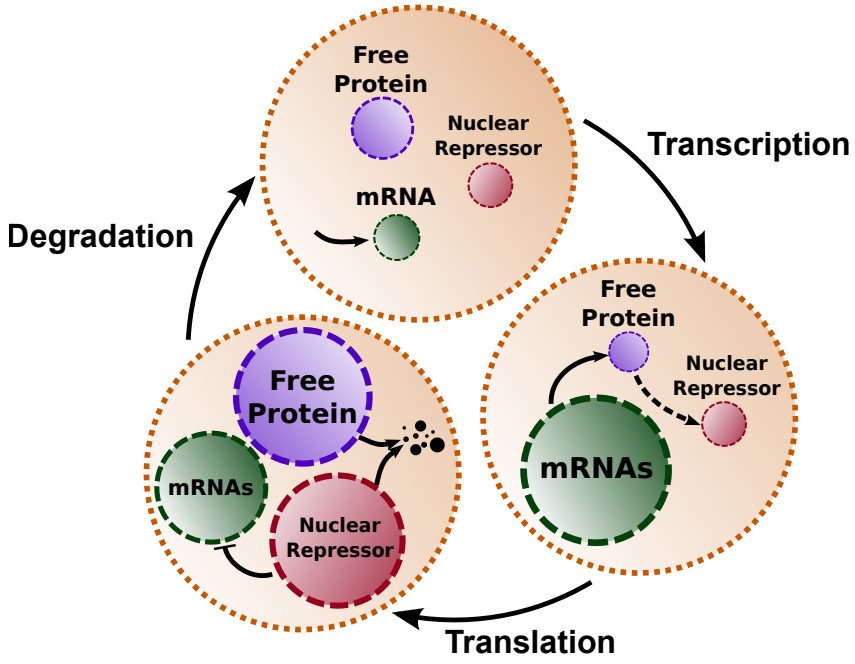


Figure 1.3: Key time delays. Negative feedback may give rise to sustained oscillations in the presence of time delays. In the circadian system, these time delays are transcription, translation, and protein degradation. Each of these rates is a function of the external environment, resulting in entrainable oscillations.

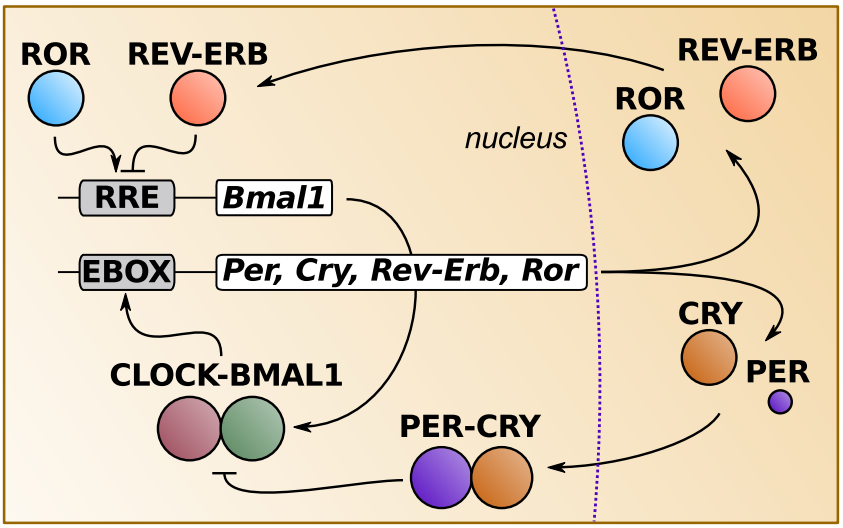


Figure 1.4: Core components of the mammalian circadian feedback loop. The mammalian feedback loop is comprised of interlocking positive and negative feedback loops.

1.1.4 Experimental techniques

While the focus of this thesis is primarily computational, a brief overview of the relevant experimental methods for studying circadian rhythms is given. Much of the work presented in this thesis has been performed in collaboration with experimental groups, including Steve Kay's group at the University of Southern California (previously UC San Diego) and Andrew Liu's group at the University of Memphis.

Circadian rhythms are studied in many different types of model organisms, and each organism requires different techniques for studying their behavior. Some of the earliest work that determined the underlying transcriptional oscillator was performed in *Drosophila*, where it was shown that mutant fly strains had altered reproduction schedules in constant dark conditions [19]. Since many analogs of the genes and proteins responsible for generating circadian rhythms in flies were also present in mammals, including *Clock, Period,* and *Cryptochrome*, much of the knowledge gained from studying fly clocks were useful in understanding oscillations in mice [19].

Early experiments determined the roles and importance of mammalian circadian genes using activity data of mouse knockouts in constant darkness [20]. For instance, the period-determining effect of CRY1 and CRY2 were found by analyzing wheel-running behavior of mice in plots known as actograms [21], as shown in figure 1.5.

While activity-level knockout data is still used to characterize the effects of circadian perturbations, such methods are costly to implement and not amenable to high-throughput experimentation. The development of luciferase reporter cell lines has rapidly changed the way circadian genes are screened [10]. In reporter cells, the coupling of a luciferase protein to a core-clock gene allows the phase of rhythms to be determined by measuring the luminescence output from cultured cells. Such systems have allowed the development of high-throughput methods for studying how

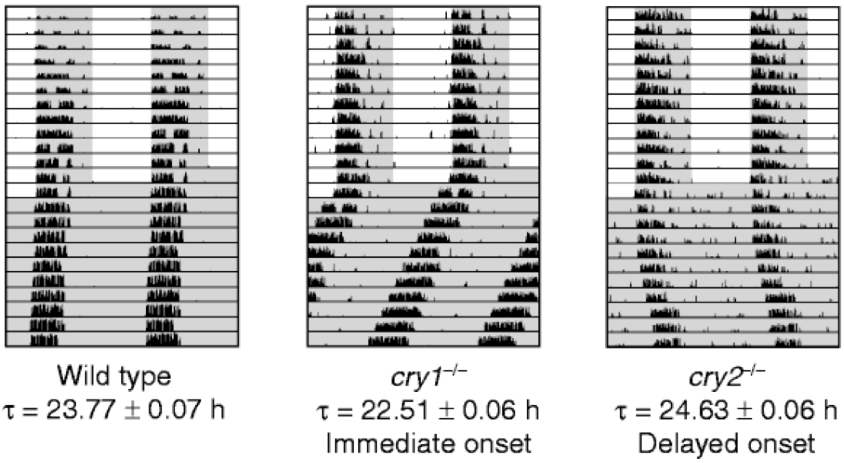


Figure 1.5: Period effects of CRY knockouts. Actograms plot the intensity of mouse wheel-running activity. The x-axis plots time of day (data is double-plotted), while the y-axis shows each subsequent day. Black bars for each day indicate the intensity of wheel-running activity. After the mice are transferred to constant darkness, the free-running period can be inferred by fitting a line to the offset in activity onset each day. Figure reprinted from [21].

siRNA [22] and small molecules [23] affect circadian parameters.

As bioluminescence experiments measure the combined output of thousands of individual cells, the synchrony of the population determines in large part the amplitude of the observed rhythms. While cells in the SCN communicate using intercellular signals to maintain synchrony in rhythms, cultured cells are suspected to lack a mechanism to spontaneously synchronize [24]. As a result, experiments will determine the rhythmicity of cultured cells by synchronizing the cells at the start of recording and recording the damped oscillations as the cells dephase [25]. While small molecule perturbations are sometimes used, the rhythms of cultured cells are acutely sensitive to their environment, and therefore simply changing the medium is often enough to induce temporary synchrony in oscillations [26].

1.2 Mathematical modeling

Circadian rhythms, due to their dynamic nature, have long been the subject of mathematical inquiry [27]. In this section, we review some of the methods previously applied to understanding circadian rhythms. While a wide range of mathematical approaches have been used to model circadian rhythms *in silico*, here we focus on techniques from ordinary differential equations (ODEs) and stochastic biochemical systems.

1.2.1 Introduction to modeling gene regulation

Gene regulatory networks are typically modeled as dynamic systems, in which rates of production and degradation of each species is explicitly considered. In figure 1.6, an example gene regulatory network is shown, in which a transcription factor activates the production of Gene 1. The genes mRNA is then translated to produce a protein product, which may or may not undergo posttranslational modifications or subcellular transport. The completed protein then modulates the production of a second mRNA for Gene 2. In this example, the production and degradation rates of the mRNA and each protein state would be explicitly modeled.

If the values of the kinetic parameters (i.e., transcription rate) were known, the model could subsequently be simulated by casting the equations as either a system of ordinary differential equations (as in section 1.2.2), or by using a chemical master equation (as in section 1.2.5). Resulting trajectories would resemble those shown in figure 1.7.

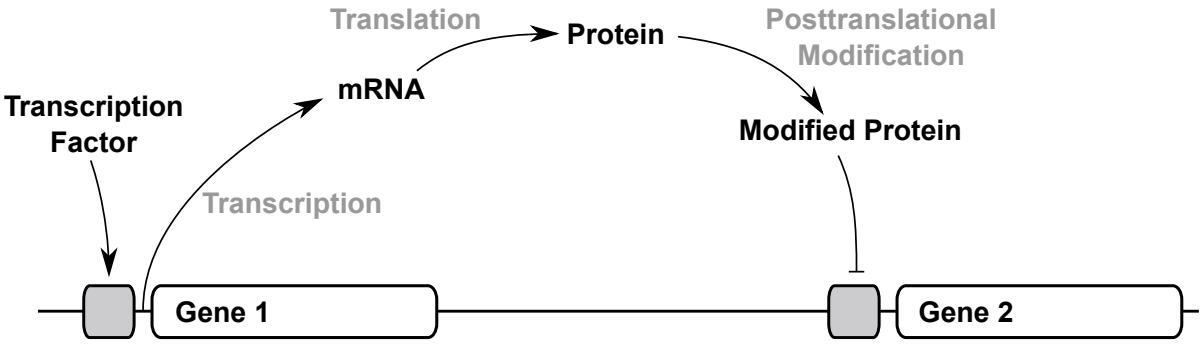


Figure 1.6: Gene regulatory networks can be modeled as dynamic systems. Transcription, translation, degradation steps are treated as chemical reactions, in which the reaction rate depends only on the current concentration of other species and various kinetic parameters.

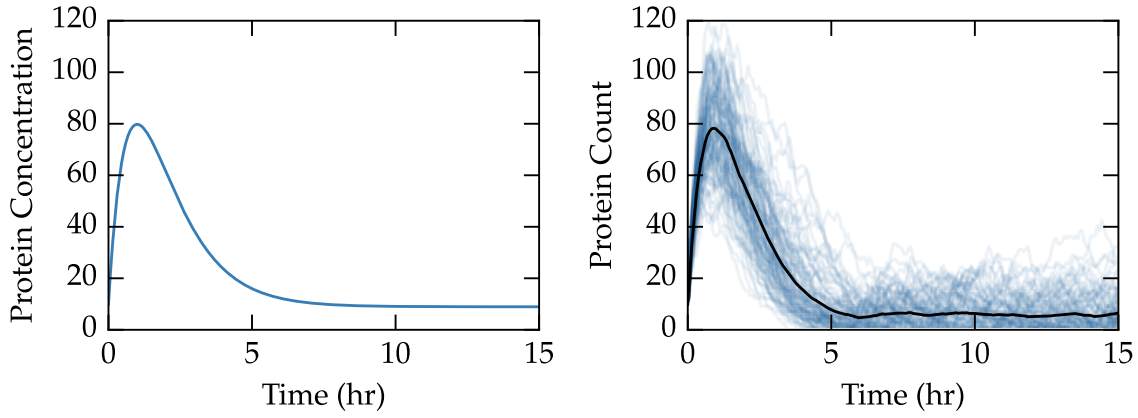


Figure 1.7: Simulated time-dependent concentration trajectories. (Left) Deterministic simulations of the dynamic system result in expected values for each state variable over time. (Right) Stochastic simulations account for the intrinsic noise present due to the low molecule counts of each reactant. From these simulations, more realistic noise estimates can be obtained at the expense of a greater computational complexity.

1.2.2 Ordinary differential equations

Ordinary differential equation (ODE) models take the general form

$$\frac{dx}{dt} = f(x(t), p), \tag{1.1}$$

in which $x(t)$ represents the concentrations of the state variables, such as mRNA and protein concentrations, f contains information on the production, degradation, and reactivity of the states, and p are the kinetic parameters which govern reaction kinetics. Limit cycle models are ODE models in which the solution approaches a steady state oscillatory trajectory, satisfying the equation:

$$\lim_{t\to\infty} [x(t+T) - x(t)] = 0. \tag{1.2}$$

The period is the smallest $T > 0$ for which equation (1.2) holds. An example periodic limit cycle is shown in figure 1.8, which plots the deterministic solution to model 6.

The points on the stable limit cycle are denoted by $x^\gamma(\theta)$ with each point assigned to a value of a phase variable $\theta \in [0, 2\pi)$. For convenience, time in equation (1.1) can be rescaled such that the period is 2π:

$$\tilde{t} = \frac{2\pi}{T}t; \quad \tilde{f} = \frac{T}{2\pi}\, f; \quad \frac{dx}{d\tilde{t}} = \tilde{f}(x(\tilde{t}), p). \tag{1.3}$$

The phase variable θ is therefore defined on the limit cycle as $\theta = \tilde{t} \mod 2\pi$, with $\theta = 0$ assigned to a unique and identifiable point. An example mapping of the phase variable θ for model 6 is shown figure 1.9

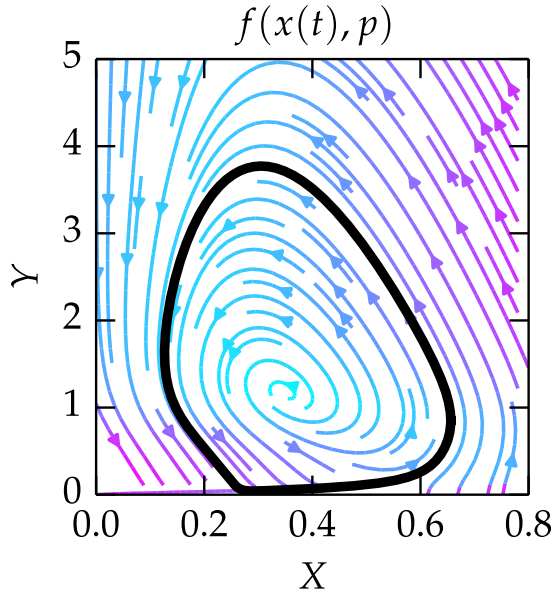


Figure 1.8: State-space representation of a deterministic limit cycle. A two-dimensional oscillator, model 6, allows the system dynamics to be easily plotted. The steady state limit cycle, $x^{\gamma}(\theta)$, is shown in dark black. Points which lie off the limit cycle will eventually converge, following the vector field arrows. Colors represent the speed with which the solution will travel, with warmer colors indicating faster speeds.

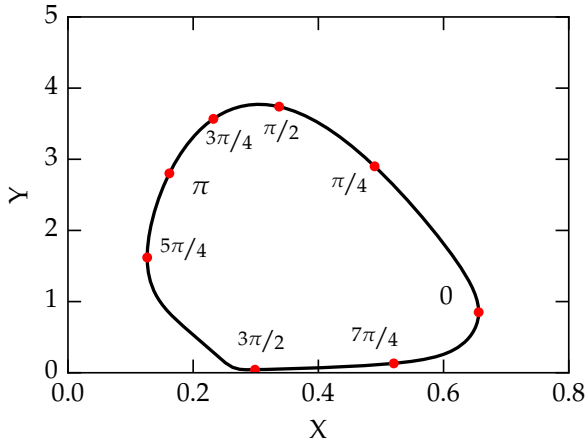


Figure 1.9: Definition of the phase variable. A phase variable θ maps each point on the limit cycle, $x^{\gamma}(t)$, to a unique value of the phase. Because the definition of θ depends on time, the distance in state-space between each phase point is not equal.

1.2.3 ODE sensitivity analysis

When analyzing ODE models, it is often important to know how the solution varies with respect to initial values or parameters. The sensitivity matrix is defined as

$$S(t) = \frac{dx(t)}{dx(0)} = \lim_{\Delta x(0) \to 0} \frac{\Delta x(t)}{\Delta x(0)}. \tag{1.4}$$

While it is possible to obtain estimates of these derivatives through finite difference methods, more accurate and less computationally intensive representations can be obtained by integrating these sensitivities along with the original ODE system [28]. The differential equation for the sensitivity system is shown in equation 1.4.

$$\frac{d}{dt}S(t) = \frac{df(x(t), p)}{dx}\, S(t) \tag{1.5}$$

Since the jacobian matrix $\frac{\partial f}{\partial y}$ is the same as that of the original ODE system, efficient implementations of the following ODEs have been developed [29].

Period sensitivity

An extension of sensitivity analysis for limit-cycle specific models raises additional complications. For instance, finding the derivative of the period with respect to parameter values requires a clever decomposition of the sensitivity matrices [30]. Period sensitivities, $\frac{\partial T}{\partial p}$, can be obtained directly from sensitivities integrated for one pass of the limit cycle (derived in [31]) through a linear solve:

$$\begin{bmatrix} \mathbf{M} - \mathbf{I} & \dot{x}(T) \\ \frac{\partial f_0}{\partial x}(x(0)) & 0 \end{bmatrix} \begin{bmatrix} \vdots \\ \frac{\partial T}{\partial p} \end{bmatrix} = \begin{bmatrix} -\mathbf{S}(T) \\ -\frac{\partial \mathbf{f}}{\partial p}(x(0)) \end{bmatrix} \tag{1.6}$$

in which **M** is the Monodromy matrix, **I** is the identity matrix, and the unknown vector contains the relevant period sensitivities.

Since parameter values often span several orders of magnitude, an often more useful measure is the relative period sensitivity, which is independent of the magnitude of the period or parameter value.

$$\frac{\partial \ln T}{\partial \ln p} = \frac{p}{T}\frac{\partial T}{\partial p} = \frac{\partial T}{T} \Big/ \frac{\partial p}{p} \tag{1.7}$$

Thus, a relative period sensitivity of 1 indicates that a 1% increase in the parameter value will result in a 1% increase in the period.

Phase sensitivity

Perturbations to state variables away from a limit cycle ultimately return, but with in a change of phase. The definition of phase can be extended to points outside the limit cycle, $x_0 \notin x^\gamma(\theta)$, by defining the phase of any point, $\Theta(x_0)$, as the initial phase of a point on the limit cycle to which x_0 will ultimately converge:

$$\Theta(x_0) = \arg\min_\theta \lim_{t\to\infty} \|x(t) - x^\gamma(\theta + t)\| \tag{1.8}$$

A phase response curve (PRC), which maps the change in phase resulting from the same perturbation applied at every phase, can therefore be found by calculating $\Delta\theta := \Theta(x^\gamma(\theta_0) + \Delta x(0)) - \theta_0$ for each θ_0. Infinitesimal PRCs, the derivative of the phase change with respect to the perturbation, are defined for state and parameter-

impulse perturbations as [32]:

$$\frac{d\theta}{dx} := \lim_{\Delta x(0)\to 0} \frac{\Delta\theta}{\Delta x(0)} \tag{1.9}$$

$$\frac{d}{d\hat{t}}\frac{d\theta}{dp} := \lim_{d,\,\Delta p\to 0} \frac{\Delta\theta}{d\ \Delta p} \tag{1.10}$$

Methods for efficiently calculating these quantities using ODE sensitivity analysis have been developed [32], with the important result that, in the limit as $d, \Delta p \to 0$,

$$\frac{d}{d\hat{t}}\frac{d\theta}{dp} = \frac{d\theta}{dx}\frac{d\hat{f}}{dp} \tag{1.11}$$

1.2.4 Common kinetic assumptions

In constructing an ODE model of a genetic regulatory network, several common kinetic motifs are employed to capture various types of activation, repression, degradation, and complex formation. In this section, we provide an overview of the kinetic assumptions used for model construction in subsequent chapters.

Mass action law

First developed approximately 150 years ago [33], the law of mass action is used for all types of chemical systems and forms the basis for subsequent kinetic constructions. In mass action kinetics, the reaction rate is proportional to the product of the activities (i.e. concentration) of each reactant. For example, in the reversible reaction

$$A + B \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} C,$$

the reaction rates for the forward and reverse reactions (r_{+1} and r_{-1}, respectively) would be

$$r_{+1} = k_{+1}[A][B]$$
$$r_{-1} = k_{-1}[C].$$

As one molecule of A and B are consumed for each molecule of C produced (and vice-versa for the reverse reaction), the ODE system describing this reaction would be:

$$\begin{aligned}
\frac{d[A]}{dt} &= k_{-1}[C] - k_{+1}[A][B] \\
\frac{d[B]}{dt} &= k_{-1}[C] - k_{+1}[A][B] \\
\frac{d[C]}{dt} &= k_{+1}[A][B] - k_{-1}[C],
\end{aligned}$$

which reaches a fixed point (equilibrium) when $d[\cdot]/dt = 0$, yielding the familiar expression for an equilibrium constant:

$$K_{eq} = \frac{k_{+1}}{k_{-1}} = \frac{[C]}{[A][B]}.$$

Michaelis-Menten Kinetics

The Michaelis-Menten rate law, first published in 1913 [34], is a commonly used assumption to simplify the kinetics of an enzyme-mediated process. The full system considered by the Michaelis-Menten model involves the reversible binding of an enzyme to a substrate, and the subsequent irreversible reaction of the enzyme-substrate

complex, forming the product:

$$E + S \underset{k_r}{\overset{k_f}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P. \tag{1.12}$$

The full ODE system for equation (1.12) is therefore

$$\frac{d[S]}{dt} = -k_f[E][S] \tag{1.13}$$
$$\frac{d[E]}{dt} = k_{cat}[ES] + k_r[ES] - k_f[E][S] \tag{1.14}$$
$$\frac{d[ES]}{dt} = k_f[E][S] - k_r[ES] - k_{cat}[ES] \tag{1.15}$$
$$\frac{d[P]}{dt} = k_{cat}[ES]. \tag{1.16}$$

Since the concentration of enzyme-substrate complex, $[ES]$ is often difficult to measure experimentally, the above system of equations is typically not applied to real data. Instead, a pseudo-steady state assumption (PSSA) is made, in which the equilibrium $E + S \underset{k_r}{\overset{k_f}{\rightleftharpoons}} ES$ is assumed to equilibrate much faster than the product formation, $ES \xrightarrow{k_{cat}} E + P$. This assumption results in $d[ES]/dt \approx 0$, which can be used to simplify the above set of equations. We also assume a conservation of enzyme mass, $[E_T] = [E] + [ES]$. By solving the now-algebraic equations 1.14-1.15 for $[E]$ and $[ES]$ and substituting the resulting expression into the remaining equations, we obtain:

$$[ES] = \frac{[E_T][S]}{k_r/k_f + [S]}$$

from which the production rate of the product is determined:

$$\frac{d[P]}{dt} = \frac{V_{max}[S]}{K_M + [S]}, \qquad V_{max} = k_{cat}E_T, \quad K_M = \frac{k_r}{k_f}.$$

While the Michaelis-Menten model is widely used for enzyme-catalyzed reactions in biology, some articles have argued that the assumptions are likely not valid for protein-protein interactions, in which the substrate and enzyme are present in comparable concentrations [35].

Enzyme inhibition

The Michaelis-Menten kinetic assumption can be extended to model the effects of enzyme inhibitors. Different types of enzyme inhibition are possible, in which an inhibitor I binds to the free enzyme E and/or the enzyme-substrate complex. In competitive inhibition, the inhibitor binds only to the free enzyme complex, blocking the binding of the substrate.

$$E + I \underset{k_{-i}}{\overset{k_i}{\rightleftharpoons}} EI \tag{1.17}$$

Adding equation (1.17) to equation (1.12) and invoking the PSSA once more, we find the final rate for the conversion of substrate to product:

$$\frac{d[P]}{dt} = \frac{V_{max}[S]}{K_M\left(1 + \frac{[I]}{K_I}\right) + [S]}; \qquad K_I = \frac{k_{-i}}{k_i}. \tag{1.18}$$

Hill-type kinetics

In biochemical systems, cooperative binding between species is often required to initiate a reaction. The kinetics of such systems follow a sigmoidal profile, in which a critical amount of reactant species is required before significant product formation is seen, but which quickly saturates to a maximum value. The equation describing cooperative binding behavior was developed in 1910 [36]. Assuming the conversion

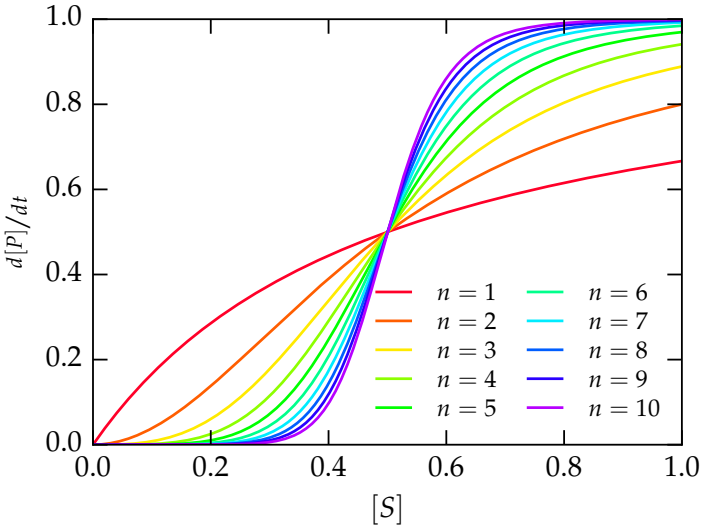


Figure 1.10: Example profiles of the hill equation for various degrees of cooperativity. The hill equation describes the kinetics of a cooperative reaction process, as shown in equation (1.19). Here $K_A = 0.5$, $V_{max} = 1$.

of reactant S to product P, the rate of production of P would follow

$$\frac{d[P]}{dt} = \frac{V_{max}[S]^n}{K_A^n + [S]^n} \tag{1.19}$$

Depending on the degree of cooperativity n, the sigmoidal profile can vary between a standard Michaelis-Menten profile (no cooperativity) to an almost digital switch, as shown in figure 1.10. Hill-type kinetics can also include enzyme inhibition, in which the K_A parameter would be scaled by inhibitor concentration in a manner similar to equation (1.18).

1.2.5 Stochastic methods

In modeling a system of chemical reactions using ODE methods, the assumption is made that the concentration of reacting species can take any continuous value. While this assumption is often appropriate for reactions in large, well-mixed vessels, it is

often not the case in reactions that take place in the cell. In mammalian cells, the median number of mRNA transcripts per cell is 17, while the median number of protein copies is $50,000$ [37]. As a result, low concentrations of state variables are often best represented by discrete steps, in which the probabilistic nature of reaction events plays a large role in determining the system dynamics.

Stochastic systems are specified by a state vector of discrete populations $x = \{x_1, \ldots, x_N\}$ and reaction vector $R = \{R_1, \ldots, R_M\}$. For each of M reactions, a stoichiometric vector v_i describes the states which are produced and consumed by each reaction, and a reaction propensity function a_i describes the likelihood of the reaction in for time step [38]. Together, these equations form the chemical master equation, a Markov (memoryless) process which can be efficiently simulated using a variety of algorithms [39].

1.2.6 Parameter estimation

Models are typically used to validate that our assumed mechanisms are mathematically feasible, and to systematically explore consequences of our kinetic assumptions. When these models differ from experimental results, we must re-analyze which of our assumptions might be incorrect and design new experiments to test these assumptions. While biological insight is typically sufficient to construct candidate models $f(x, p)$ using standard assumptions outlined in section 1.2.4, values for the kinetic parameters are typically derived by fitting experimental data. In this section, we describe optimization techniques which have proven successful in finding descriptive parameter sets for models of circadian rhythms.

Cost function

The first step in finding a suitable parameter set is defining a cost function, $C(p)$, which ascribes a value to how well a particular parameter set fits the data. By convention, the optimum parameter set is that which minimizes $C(p)$. While for many biological systems a suitable cost function is the squared error of the model, a cost function can also include information on model derivatives, oscillatory period, and timing of various events (i.e., peak gene expression). Cost functions of limit cycle models pose a particularly difficult optimization challenge, as a limit cycle does not always exist for any given parameter set. Thus, the fitness landscape for most parameter space is very flat where limit cycle solutions do not exist, with isolated regions of where limit cycle oscillations are possible. Additionally, solving for the limit cycle solution given a parameter set is a nested optimization problem, resulting in long computation times. For these reasons, stochastic global optimization procedures are often employed which efficiently search high dimensional parameter space without the need for derivative information of $C(p)$. An overview of the computational approach typically employed for such systems is given in figure 1.11.

Global optimization

In this thesis, global optimization of model parameters is typically achieved via an evolutionary algorithm approach [40]. An evolutionary algorithm iteratively improves a parameter guess in an easily-parallelized fashion by mimicking natural evolution. The steps involved are as follows, shown schematically in figure 1.12.

1. **Initialization** An initial population of parameter guesses are generated from all parameter values, typically by a random number generation.

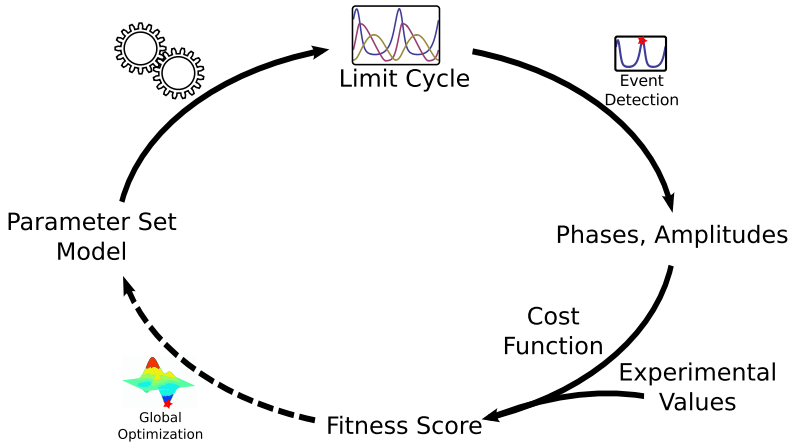


Figure 1.11: An overview of the nested optimization of fitting limit cycle models. Given a model and parameter set guess, a limit cycle is calculated via a single-shooting approach. A cost function then scores the timing of peak mRNA and protein species against known experimental values. Given the fitness score of many parameter guesses, a global optimization algorithm iteratively improves the parameter guess to minimize the cost function.

2. **Selection** The fitnesses (inverse of the cost) of the current population are evaluated, and a subsample of the original population is selected. A selection algorithm must randomly select from the current population, with some degree bias given towards individuals with the highest fitness.

3. **Crossover** Each member of the subsequent generation of solutions is generated by combining features from two (or more) of the selected individuals in the previous generation. For a problem consisting of many individual parameters, a common choice is to randomly select each parameter from one parent or the other.

4. **Mutation** The final step involves adding a small chance to change a parameter in the child parameter set. For continuous parameter values, this is often implemented as a small chance to add normally-distrusted error to one or more of the parameter values.

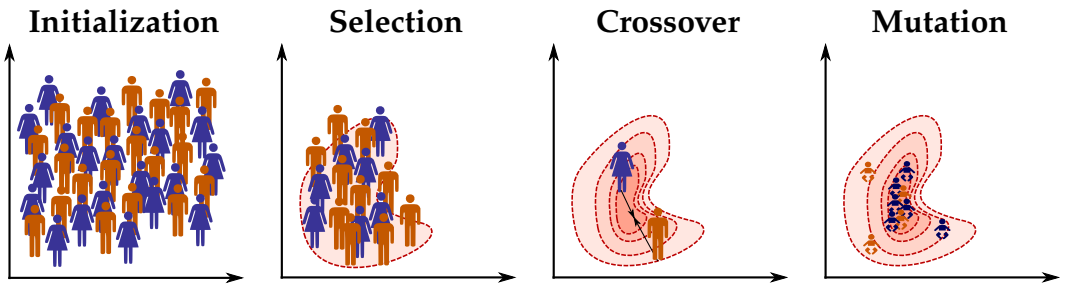


Figure 1.12: An overview of global optimization via evolutionary strategy. A schematic of a two-dimensional parameter space is shown. Solutions are initially distributed randomly across the parameter landscape, but solutions with a high fitness are preferentially kept. By combining and slightly altering parameter values from existing solutions, each generation iteratively approaches the highest fitness.

1.2.7 Previous models of circadian rhythms

Since mathematical methods have been used to study circadian rhythms for some time, many existing models of circadian gene regulation exist. A summary of the models discussed in this section is presented in table 1.1. Early examples include Goodwin's 1965 model of enzymatic oscillations, demonstrating that nonlinear oscillations could capture many of the features of biological rhythms [41]. Later models sought to describe specific proteins in the circadian cycle. Goldbeter's 1995 model described one gene in the circadian clock of *Drosophila*, and demonstrated that a negative feedback loop with sufficient nonlinearity and time delay was required for circadian oscillations [42].

Subsequent models have attempted to generate a more complete representation of the circadian network. Leloup and Goldbeter's 2003 model of the mammalian circadian clock considers both the positive and negative regulatory loops [2], demonstrating the anti-phase oscillations of the positive and negative elements. Forger and Peskin's 73-state model, and its analogous stochastic version, provides an example of the complexity of the known post-translational regulation of core clock components

First Author	Year	States	Parameters	Citation
Goodwin	1965	2	5	[41]
Goldbeter	1995	5	15	[42]
Leloup	2003	16	52	[2]
Forger	2003	73	38	[43]
Mirsky	2009	21	132	[44]
Kim	2012	180	70	[45]

Table 1.1: Summary of ODE models for the circadian network. While Goodwin's model considers an arbitrary oscillatory biochemical system, it has been frequently invoked to model circadian oscillations in a simple framework.

[43, 46].

More recently, models have focused on explaining new theories on how circadian rhythms are organized, rather than attempting to providing a complete *in silico* representation of current knowledge. A model by To *et al.,* 2007 coupled together cells modeled by Leloup & Goldbeter's 2003 model to demonstrate how heterogeneous populations of oscillators could synchronize in an SCN environment using the neurotransmitter vasoactive intestinal polypeptide (VIP) [47]. Following the differentiation between the properties of single-cell oscillators and populations, Mirsky *et al.,*, 2009 developed a model to show how single-cell knockout phenotypes could be differentiated from those at the population-level [44]. Relógio *et al.,* 2011 demonstrated that positive feedback through ROR/REV-ERB had a stabilizing effect on the clock [1], while a model from Kim & Forger, 2012 demonstrated the importance of stoichiometry between positive and negative components [45]. However, despite the wide availability of mathematical models for circadian rhythms, demonstrating that new theories are mathematically rigorous often requires the creation of a new model.

1.3 Thesis Overview

In chapter 2, a new model of the core circadian feedback loop is developed in order to explain experimental observations on the effect of a new small molecule modulator. This new model enables new insights into the functional roles of the two isoforms of mammalian Cryptochrome, CRY1 and CRY2. In chapter 3, a method of fitting oscillatory models is presented which uses nonlinear programming rather than the more computationally intensive global optimization strategies typically employed with circadian models. Using such a method, a bootstrap approach is possible which allows the quantification of confidence intervals in the predicted responses generated from the models. This method is then applied to a set of three circadian models in chapter 4 to explain the separate amplitude effects of two small molecule regulators by finding predictions which are consistent across slight differences in kinetic assumptions and parameter value. In chapter 5, we derive a new type of sensitivity analysis for oscillatory models that we denote the amplitude response curve (ARC). The ARC allows the effect from a temporary pulse, coming perhaps from a small molecule therapeutic, on amplitude to be calculated. This method allows the systematic investigation of factors which change clock amplitude to proceed efficiently, without the need for more expensive stochastic simulations of a population of oscillators. Finally, in chapter 6, it is shown that damping rate of cultured circadian reporter cells can be a reliable indicator of single-cell noise. Using this insight, we analyze data from high-throughput siRNA screens to determine the effects of genome-wide gene knockdowns on the stochastic noise. From this analysis we conclude that the unperturbed clock is highly optimized for robustness, as it lies near the pareto front of high amplitude and low stochastic noise.

Chapter 2

Modeling the core circadian feedback loop to determine the effect of KL001[1]

2.1 Background

2.1.1 Small molecule modulators

Circadian rhythms help to regulate metabolic homeostasis, and therefore compromised circadian oscillations often result in negative health consequences. Since circadian rhythms may be damped by lifestyle factors, many researchers have begun searching for pharmacological agents which might alter circadian rhythmicity. Several proof-of-concept molecules have been found which cause dose-dependent period change in cultured fibroblasts [48]. While chemical biology approaches are able to localize the effects of these molecules to specific targets in the circadian network, it is often difficult to reconcile individual rate changes with the observed systems-

[1] Portions of this chapter are published in T. Hirota, J. W. Lee, P. C. St. John, M. Sawa, K. Iwaisako, T. Noguchi, P. Y. Pongsawakul, T. Sonntag, D. K. Welsh, D. A. Brenner, F. J. Doyle III, P. G. Schultz, and S. A. Kay, "Identification of Small Molecule Activators of Cryptochrome.," *Science*, vol. 337, pp. 1094–1097, July 2012.

level changes in period or amplitude. Mathematical models are therefore essential to understanding the effects of such drugs, as they are able to verify that hypothesized mechanisms of action are mathematically consistent. Furthermore, mathematical models provide a systematic framework by which to examine consequences of various mechanistic assumptions - many of which can be subsequently verified by experimental inquiry. The search for small molecule modulators of circadian rhythms therefore yields the added benefit of strengthening our knowledge of the core clock feedback circuit, as inconsistencies between assumed kinetics and experimental results often lead to a more refined knowledge of the network's dynamics.

This chapter describes a collaborative project in which experimental results inspired the creation of a new model of the core mammalian feedback network. A new small molecule modulator of Cryptochrome, KL001, was identified through forward chemical genetic screening [49]. The compound was found to simultaneously increase the half-life of CRY and cause period lengthening. In order to explain these new experimental results while maintaining consistency with published evidence, a simple mathematical model of the PER/CRY negative feedback loop was constructed.

The functional roles of the cryptochrome isoforms (*Cry1* and *Cry2*) in circadian rhythms have long eluded biologists, as although they share a similar structure, perturbations to these genes typically result in opposite trends [50]. Existing genetic evidence and siRNA knockdown studies agree that while suppressing *Cry1* leads to shorter periods, suppressing *Cry2* leads to longer periods [21, 22]. However, some evidence shows the similarity between the two *Crys*. The knockouts[2] $Cry1^{+/-}$ $Cry2^{-/-}$ and $Cry1^{-/-}$ $Cry2^{+/-}$ both show shorter wheel-running activity their re-

[2]Wild type mammals, denoted with the "$+/+$" superscript, receive two copies of every gene from each parent. Homozygous knockouts, denoted "$-/-$", have no functional copies of the gene. Heterozygous knockouts, denoted "$+/-$", receive only one functional copy and therefore express the gene at lower levels.

spective WT/double knockouts *Cry1*$^{+/+}$ *Cry2*$^{-/-}$ and *Cry1*$^{-/-}$ *Cry2*$^{+/+}$ [21]. Using new data from the application of KL001 together with existing experimental results, a mechanism was hypothesized to reconcile these seemingly contradictory results. The experimentally motivated assumptions underlying this hypothesized mechanism are described in the following sections, with descriptions of how they are supported by the literature.

Assumptions on network structure

In the core feedback loop, *Per* and *Cry* transcription is activated by the CLOCK-BMAL1 complex. The protein products of these genes are transported to the nucleus after a time delay and bind to CLOCK-BMAL1, repressing their own transcription. Evidence indicates that an interaction of PER and CRY proteins is required for the timely nuclear entry of both proteins [16, 51, 52]. PER and CRY bind and enter the nucleus in a 1:1 stoichiometric ratio, similar to more complicated models of mammalian circadian rhythms [2, 43, 44].

Assumptions on protein stoichiometry

Quantitative protein assays have revealed that the PER proteins are present at levels nearly ten times less than the CRYs, and are completely consumed at their trough. The CRYs, on the other hand, only fall to about one quarter of their peak value [17]. These results indicate that PER is the rate-limiting component for nuclear entry. This configuration explains why, since *Cry* is in excess, a knockout of either *Cry* alone does not lead to loss of function at the tissue level. Additionally CRY1, the more dominant repressor, is consistently found at higher levels than CRY2. However, since the ratio of nuclear to total protein is nearly equivalent for CRY1 and CRY2, PER must import each with equal affinity [17, 53].

Assumptions on protein degradation

Studies of the simplest cellular oscillators have illustrated the importance of the degradation of repressor complexes in determining the period of oscillation [54]. In the mammalian circadian clock, FBXL3 has been indicated as a key protein in clearing the nucleus of CRY, and siRNA knockdowns of *Fbxl3* result in period lengthening [22]. While FBXL3 localizes primarily in the nucleus, CRY2 possesses a distinct phosphorylation domain that flags it for cytosolic proteolysis [55]. Knocking down this domain shortens the period, indicating that degradation rates play a major role in determining the oscillatory period and that stabilization in the cytosol and nucleus are likely not equivalent.

Assumptions on time delays

Similar to *Drosophila*, mammalian circadian oscillations are generated by three main time delays in the negative feedback loop: mRNA transcription, nuclear localization of protein, and degradation of the nuclear complexes [56]. In order to capture these dynamics without introducing unnecessary stiffness, only the concentrations of nuclear mRNA, cytosolic unbound protein, and nuclear protein complexes are explicitly considered.

2.2 Results

2.2.1 A new model for period regulation

Since the circadian clock is largely insensitive to the total amount *Cry*, it follows that the core negative loop is comprised of two redundant coupled inhibitory mechanisms. The natural periods of these two loops (running in isolation of the other) can

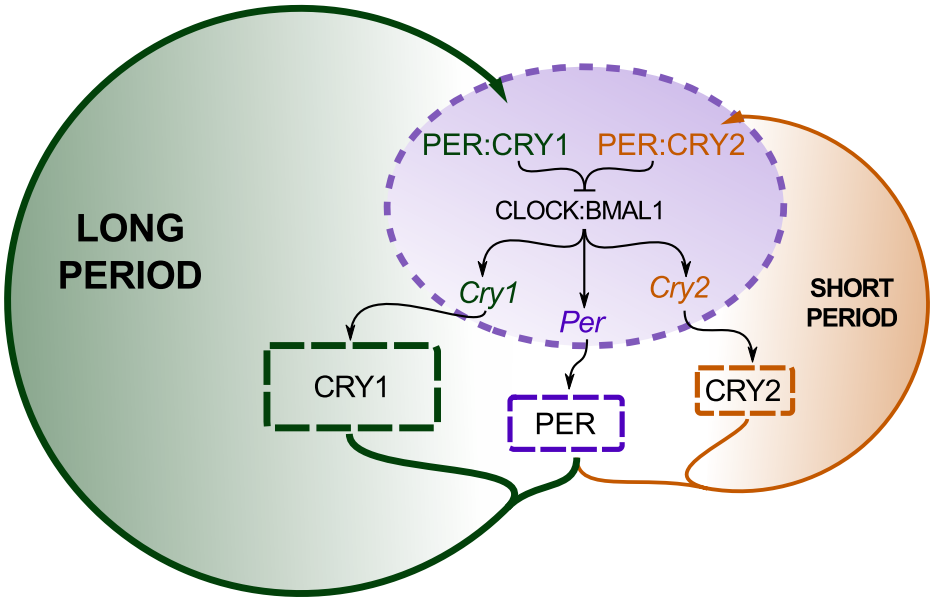


Figure 2.1: Connectivity of model 1. Schematic for the core negative feedback loop, consisting of the two redundant *Cry* mechanisms. The wild-type period, consisting of both loops, is a balance between the short and long oscillations. Figure reprinted from [49].

be inferred from their corresponding knockout phenotype, i.e., the CRY1 feedback loop has a long period, while the CRY2 loop has a short period. In the wild-type phenotype, the two isoforms of CRY must compete for the available PER, and thus the nuclear CRY1/CRY2 levels are constrained with a ratio proportional to their relative expression. This suggests that the period of the clock is governed by the nuclear CRY1/CRY2 ratio, where higher amounts of *Cry1* shift the clock closer to the $Cry2^{-/-}$ phenotype, and higher amounts of *Cry2* shift the clock closer to the $Cry1^{-/-}$ phenotype, described in figure 2.1.

The CRY1/CRY2 ratio

In silico modeling was used to test the feasibility of the above hypothesis, and to investigate possible mechanisms by which the CRY1/CRY2 ratio might control the pe-

riod of oscillations. The model uses eight state variables for the three mRNA species (*Per*, *Cry1*, and *Cry2*), three cytosolic proteins, and two nuclear proteins. The differential equations for each state were formulated using standard Hill-type repression, Michaelis-Menten, and mass action kinetics. The 21 unknown kinetic parameters were found by fitting the stoichiometric data from [17] using a genetic algorithm approach, requiring correct periods for the $Cry1^{-/-}$ and $Cry2^{-/-}$ knockouts.

We investigated two possible mechanisms to explain the long/short period phenotype of the CRY1/CRY2 loops. First, CRY1 and CRY2 were allowed to have different inhibitory efficiency, as demonstrated experimentally [57]. However, optimizations with this structure were unable to find a parameter set with appropriate period sensitivities, described in section 2.2.2. Instead, we investigated whether the difference in potency between *Crys* could be explained through difference in degradation rates. To this end, the degradation rates of nuclear CRY1 and CRY2 were allowed to differ, while their inhibition constants were set to be equal. Using this configuration, the structure was able to fit the experimental sensitivities. In the optimized parameter set, the degradation rate of nuclear CRY2 is higher than that of CRY1, such that for a constant total nuclear CRY, higher fractions of CRY2 cause the repressive complexes to be cleared faster. Since PER is the limiting reagent in nuclear entry, the amount of total CRY that enters the nucleus is largely insensitive to perturbations in cytosolic CRY expression. In effect, the two isoforms of CRY must compete for the available PER, and thus the nuclear CRY1/CRY2 levels are constrained with a ratio proportional to their relative expression. The period of the clock is thus governed by the nuclear CRY1/CRY2 ratio, where higher amounts of *Cry1* shift the clock closer to the $Cry2^{-/-}$ phenotype, and higher amounts of *Cry2* shift the clock closer to the $Cry1^{-/-}$ phenotype. Figure 2.2 shows the time varying total complex concentration under various perturbations, with the relative contributions from CRY1 and CRY2

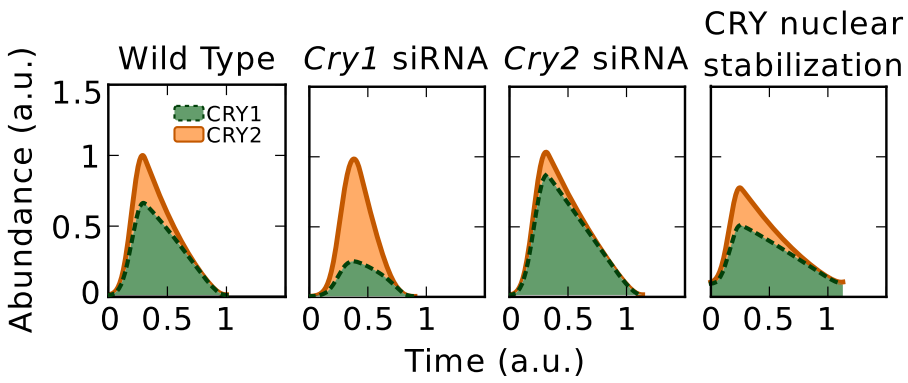


Figure 2.2: Repressor degradation profiles. Time course plots of the simulated total nuclear repressor concentration for the degradation model under various clock conditions. The relative contribution of CRY1 and CRY2 are shaded green and orange, respectively. Longer periods occur when a higher percentage of the nuclear complex is CRY1, taking longer to degrade. Figure reprinted from [58].

highlighted.

Model equations

The model is formulated as set of 8 ordinary differential equations shown in model 1 (for the degradation-based model) and model 2 (for the activation-based model). Additional assumptions made while formulating the model equations are listed below:

- For simplicity and ease of parameter estimation, the model only considers the three genes *Per*, *Cry1*, and *Cry2*, not explicitly considering the three known isoforms of *Per*.

- EBOX activators *Clock* and *Bmal1* are considered constitutively expressed and are represented by the v_{txn} parameters.

- Repression of CLOCK-BMAL1 activity is attained through Hill-type inhibition. The Hill coefficient is fixed at 3, which was found to provide sufficient non-

linearity for oscillations. See below for the derivation of the repressor input function used in model 1 and model 2.

- The degradation of the mRNA species and cytoplasmic proteins is assumed to follow standard Michaelis-Menten kinetics.

- The degradation kinetics of the nuclear proteins are assumed to follow Michaelis-Menten kinetics. Because both **C1N** and **C2N** are thought to be degraded by the same pathway in the nucleus, kinetic equations using the *pseudo* steady-state hypothesis were derived for two substrates sharing the same enzyme. As a result, each CRY isoform acts as an inhibitor to the other's degradation. This mechanism is summarized in figure 2.3.

Derivation of a shared enzyme degradation rate

The rate equations associated with a shared-enzyme degradation mechanism are derived from the equilibrium relationships:

$$[\mathrm{E}] + [\mathrm{S1}] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [\mathrm{ES1}] \xrightarrow{k_{d,1}} [\mathrm{E}]$$

$$[\mathrm{E}] + [\mathrm{S2}] \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} [\mathrm{ES2}] \xrightarrow{k_{d,2}} [\mathrm{E}]$$

$$[\mathrm{E}_t] = [\mathrm{E}] + [\mathrm{ES1}] + [\mathrm{ES2}] \tag{2.9}$$

The end goal is the degradation rates of the two enzyme complexes:

$$\begin{aligned} r_{d,1} &= k_{d,1}\,[\mathrm{ES1}] \\ r_{d,2} &= k_{d,2}\,[\mathrm{ES2}] \end{aligned} \tag{2.10}$$

Model 1: A degradation-based model of the core feedback loop. Lower case letters (p: *Per*, c1: *Cry1*, c2: *Cry2*) are mRNA state variables. Upper-case letters (P: PER, C1: CRY1, C2: CRY2) are the free (cytosolic) proteins. C1N: CRY1 and C2N: CRY2 are the nuclear proteins.

$$\frac{d\mathbf{p}}{dt} = \frac{v_{\text{txn},\mathbf{p}}}{k_{\text{txn},\mathbf{p}} + (\mathbf{C1N} + \mathbf{C2N})^3} - \frac{v_{\text{deg},\mathbf{p}}\ \mathbf{p}}{k_{\text{deg},\mathbf{p}} + \mathbf{p}} \tag{2.1}$$

$$\frac{d\mathbf{c1}}{dt} = \frac{v_{\text{txn},\mathbf{c1}}}{k_{\text{txn},\mathbf{c}} + (\mathbf{C1N} + \mathbf{C2N})^3} - \frac{v_{\text{deg},\mathbf{c1}}\ \mathbf{c1}}{k_{\text{deg},\mathbf{c}} + \mathbf{c1}} \tag{2.2}$$

$$\frac{d\mathbf{c2}}{dt} = \frac{v_{\text{txn},\mathbf{c2}}}{k_{\text{txn},\mathbf{c}} + (\mathbf{C1N} + \mathbf{C2N})^3} - \frac{v_{\text{deg},\mathbf{c2}}\ \mathbf{c2}}{k_{\text{deg},\mathbf{c}} + \mathbf{c2}} \tag{2.3}$$

$$\begin{aligned}\frac{d\mathbf{P}}{dt} = k_{\text{tln},\mathbf{p}}\ \mathbf{p} - \frac{v_{\text{deg},\mathbf{P}}\ \mathbf{P}}{k_{\text{deg},\mathbf{P}} + \mathbf{P}} - v_{\text{a},\mathbf{CP}}\ \mathbf{P}\ \mathbf{C1} + v_{\text{d},\mathbf{CP}}\ \mathbf{C1N} \\ - v_{\text{a},\mathbf{CP}}\ \mathbf{P}\ \mathbf{C2} + v_{\text{d},\mathbf{CP}}\ \mathbf{C2N}\end{aligned} \tag{2.4}$$

$$\frac{d\mathbf{C1}}{dt} = \mathbf{c1} - \frac{v_{\text{deg},\mathbf{C1}}\ \mathbf{C1}}{k_{\text{deg},\mathbf{C}} + \mathbf{C1}} - v_{\text{a},\mathbf{CP}}\ \mathbf{P}\ \mathbf{C1} + v_{\text{d},\mathbf{CP}}\ \mathbf{C1N} \tag{2.5}$$

$$\frac{d\mathbf{C2}}{dt} = \mathbf{c2} - \frac{v_{\text{deg},\mathbf{C2}}\ \mathbf{C2}}{k_{\text{deg},\mathbf{C}} + \mathbf{C2}} - v_{\text{a},\mathbf{CP}}\ \mathbf{P}\ \mathbf{C2} + v_{\text{d},\mathbf{CP}}\ \mathbf{C2N} \tag{2.6}$$

$$\frac{d\mathbf{C1N}}{dt} = -\frac{v_{\text{deg},\mathbf{CP}}\ \mathbf{C1N}}{k_{\text{deg},\mathbf{CP}} + \mathbf{C1N} + \mathbf{C2N}} + v_{\text{a},\mathbf{CP}}\ \mathbf{P}\ \mathbf{C1} - v_{\text{d},\mathbf{CP}}\ \mathbf{C1N} \tag{2.7}$$

$$\frac{d\mathbf{C2N}}{dt} = -\frac{\left(v_{\text{deg},\mathbf{CP}}\ m_{\mathbf{C2N}}\right)\ \mathbf{C2N}}{k_{\text{deg},\mathbf{CP}} + \mathbf{C2N} + \mathbf{C1N}} + v_{\text{a},\mathbf{CP}}\ \mathbf{P}\ \mathbf{C2} - v_{\text{d},\mathbf{CP}}\ \mathbf{C2N} \tag{2.8}$$

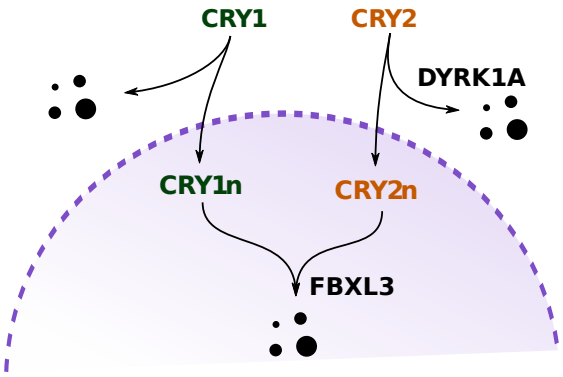


Figure 2.3: Degradation of both CRY isoforms is regulated by the same enzyme. In developing the model equations for model 1, a key assumption was that both isoforms of Cryptochrome were degraded by the same enzyme. Without such a scheme, the shortest periods were obtained by maximizing degradation through equal nuclear entry of both isoforms.

Model 2: Changed equations for the activation-based model. The remainder of the model equations (not duplicated below) are found in model 1

$$\frac{d\mathbf{p}}{dt} = \frac{v_{\text{txn},\mathbf{p}}}{k_{\text{txn},\mathbf{p}} + (m_{\mathbf{C2N}}\,\mathbf{C1N} + \mathbf{C2N})^3} - \frac{v_{\text{deg},\mathbf{p}}\,\mathbf{p}}{k_{\text{deg},\mathbf{p}} + \mathbf{p}} \tag{2.1*}$$

$$\frac{d\mathbf{c1}}{dt} = \frac{v_{\text{txn},\mathbf{c1}}}{k_{\text{txn},\mathbf{c}} + (m_{\mathbf{C2N}}\,\mathbf{C1N} + \mathbf{C2N})^3} - \frac{v_{\text{deg},\mathbf{c1}}\,\mathbf{c1}}{k_{\text{deg},\mathbf{c}} + \mathbf{c1}} \tag{2.2*}$$

$$\frac{d\mathbf{c2}}{dt} = \frac{v_{\text{txn},\mathbf{c2}}}{k_{\text{txn},\mathbf{c}} + (m_{\mathbf{C2N}}\,\mathbf{C1N} + \mathbf{C2N})^3} - \frac{v_{\text{deg},\mathbf{c2}}\,\mathbf{c2}}{k_{\text{deg},\mathbf{c}} + \mathbf{c2}} \tag{2.3*}$$

$$\frac{d\mathbf{C2N}}{dt} = -\frac{v_{\text{deg},\mathbf{CP}}\,\mathbf{C2N}}{k_{\text{deg},\mathbf{CP}} + \mathbf{C2N} + \mathbf{C1N}} + v_{\text{a},\mathbf{CP}}\,\mathbf{P}\,\mathbf{C2} - v_{\text{d},\mathbf{CP}}\,\mathbf{C2N} \tag{2.8*}$$

As was done in section 1.2.4, we invoke the standard pseudo steady-state assumption and set $d[\mathrm{ES}]/dt = 0$, and obtain the following production = consumption equalities for [ES1]:

$$k_1[\mathrm{E}][\mathrm{S}] = k_{-1}[\mathrm{ES1}] + k_{d,1}[\mathrm{ES1}] \tag{2.11}$$

Solving equation (2.11) for ES1 and substituting in equation (2.9), we obtain

$$\left(\frac{k_{-1} + k_{d,1}}{k_1}\right)[\mathrm{ES1}] = ([\mathrm{E}_t] - [\mathrm{ES1}] - [\mathrm{ES2}])\,[\mathrm{S1}]$$

After defining a useful combined rate constant,

$$K_1 \equiv \frac{k_{-1} + k_{d,1}}{k_1}$$

we can further simplify

$$\begin{aligned} K_1[\mathrm{ES1}] + [\mathrm{S1}][\mathrm{ES1}] &= ([\mathrm{E}_t] - [\mathrm{ES2}])[\mathrm{S1}] \\ [\mathrm{ES1}] &= \frac{([\mathrm{E}_t] - [\mathrm{ES2}])[\mathrm{S1}]}{K_1 + [\mathrm{S1}]} \end{aligned} \tag{2.12}$$

If the same operations are performed for [ES2] and plugged into equation (2.12).

$$\begin{aligned} [\mathrm{ES1}] &= \frac{\left([\mathrm{E}_t] - \frac{([\mathrm{E}_t] - [\mathrm{ES1}])[\mathrm{S2}]}{K_2 + [\mathrm{S2}]}\right)\mathrm{S1}]}{K_1 + [\mathrm{S1}]} \\ [\mathrm{ES1}] &= \frac{[\mathrm{E}]_t[\mathrm{S1}]}{K_1 + [\mathrm{S1}]} - \frac{[\mathrm{E}]_t[\mathrm{S1}][\mathrm{S2}]}{(K_1 + [\mathrm{S1}])(K_2 + [\mathrm{S2}])} + \frac{[\mathrm{ES1}][\mathrm{S1}][\mathrm{S2}]}{(K_1 + [\mathrm{S1}])(K_2 + [\mathrm{S2}])} \\ [\mathrm{ES1}] &= \left(\frac{[\mathrm{E}]_t[\mathrm{S1}]}{K_1 + [\mathrm{S1}]}\right)\left(\frac{1 - \frac{[\mathrm{S2}]}{K_2 + [\mathrm{S2}]}}{1 - \frac{[\mathrm{S1}][\mathrm{S2}]}{(K_1 + [\mathrm{S1}])(K_2 + [\mathrm{S2}])}}\right) \end{aligned} \tag{2.13}$$

The second fraction from equation (2.13) is then pulled into the denominator of the first and simplified further.

$$
\begin{aligned}
[\mathrm{ES1}] &= \frac{[\mathrm{E}]_t[\mathrm{S1}]}{\left(\frac{K_1+[\mathrm{S1}]-\frac{[\mathrm{S1}][\mathrm{S2}]}{K_2+[\mathrm{S2}]}}{1-\frac{[\mathrm{S2}]}{K_2+[\mathrm{S2}]}}\right)} \\
[\mathrm{ES1}] &= \frac{[\mathrm{E}]_t[\mathrm{S1}]}{\left(\frac{K_2+[\mathrm{S2}]}{K_2}\right)\left(K_1+[\mathrm{S1}]-\frac{[\mathrm{S1}][\mathrm{S2}]}{K_2+[\mathrm{S2}]}\right)} \\
[\mathrm{ES1}] &= \frac{K_2[\mathrm{E}]_t[\mathrm{S1}]}{(K_1+[\mathrm{S1}])(K_2+[\mathrm{S2}])-[\mathrm{S1}][\mathrm{S2}]} \\
[\mathrm{ES1}] &= \frac{K_2[\mathrm{E}]_t[\mathrm{S1}]}{K_1K_2+K_2[\mathrm{S1}]+K_1[\mathrm{S2}]} \\
[\mathrm{ES1}] &= \frac{K_2[\mathrm{E}]_t[\mathrm{S1}]}{K_1K_2+K_2[\mathrm{S1}]+K_1[\mathrm{S2}]}
\end{aligned}
$$

Dividing top and bottom by K2:

$$
[\mathrm{ES1}] = \frac{[\mathrm{E}]_t[\mathrm{S1}]}{K_1+[\mathrm{S1}]+\frac{K_1}{K_2}[\mathrm{S2}]} \tag{2.14}
$$

Substituting equation (2.14) into equation (2.10) and setting $K_1 = K_2$, $k_{d,1} \neq k_{d,2}$, we obtain the final shared rate laws (with an equivalent analysis for [ES2])

$$
r_{d,1} = \frac{V_{\mathrm{max},1}[\mathrm{S1}]}{K_M+[\mathrm{S1}]+[\mathrm{S2}]}
$$

$$
r_{d,2} = \frac{V_{\mathrm{max},2}[\mathrm{S2}]}{K_M+[\mathrm{S1}]+[\mathrm{S2}]}
$$

Derivation of Hill-type repression formulas

To derive the equations for the Hill-type inhibition, allowing for different repressive activities (model 2), we start with the standard equation for Hill-type regulation:

$$\frac{v_{\max}[A]^n}{K_m^n + [A]^n} \tag{2.15}$$

Allowing for competitive inhibition with two inhibitors, K_m is replaced with the apparent Michaelis-Menten constant

$$K_m^{\text{app}} = K_m \left(1 + \frac{[I_1]}{K_{i,1}} + \frac{[I_2]}{K_{i,2}}\right) \tag{2.16}$$

By assuming constitutive activator concentrations and non-dimensionalizing $[I^\star] = \frac{[I]}{K_{i,2}}$, $m = \frac{K_{i,2}}{K_{i,1}}$:

$$\frac{v_{\max^\star}}{K_m^\star + \left(m\,[I_1^\star] + [I_2^\star]\right)^n}$$

Assuming equal repressive activity ($K_{i,2} = K_{i,1}$), $m = 1$, we obtain the rate equation used in the degradation model.

2.2.2 Parameter estimation

The model equations, specifying the state and parameter dependent time derivatives of each concentration variable, were written in python using the CasADi computer algebra package [59]. The model was simulated using the SUNDIALS suite of ODE solvers [60]. A parameter-dependent cost function was developed that assigns numerical values reflecting how well a parameter set fits desired features, taken from [17]. The first step in the cost function evaluation is the numerical solution of the limit cycle, described previously [31]. If a limit cycle was unable to be found, the cost

function returns a maximum value. Otherwise, the cost function returns a squared difference from the desired value. A priority weight was also attached to each cost entry, such that more important costs would be prioritized. A description of each entry in the cost function, along with the value for experimental and final model is shown in table 2.1 (degradation-based model, model 1; activation-based model, model 2). A genetic algorithm strategy was employed to minimize the cost function, which was used previously for optimizing circadian parameters [44]. In the method, 5000 solutions are calculated at random parameter values chosen on a log scale, with $p_{init} \in (10^{-3}, 10^{1})$. Those with the best cost function scores are kept and used to generate subsequent solutions. This procedure is iterated for up to 2000 generations, or until convergence criteria are met.

Table 2.1: Summary of cost function entries. The weights for each entry were chosen based on the relative importance of the desired behavior. Entries 1-9 and 14-18 were obtained from the data presented in [17]. SiRNA sensitivities were obtained were taken from [22].

	Description	Weight	Desired [ref]	Degradation	Activation
1	*Per* mRNA peak-trough ratio $\frac{y_{\max}(Per)}{y_{\min}(Per)}$	0.5	> 20 [17]	large	large
2	*Cry1* mRNA peak-trough ratio $\frac{y_{\max}(Cry1)}{y_{\min}(Cry1)}$	0.5	2.155 [17]	8.942	3.748

	Description	Weight	Desired [ref]	Degradation	Activation
3	*Cry2* mRNA peak-trough ratio $\frac{y_{\max}(Cry2)}{y_{\min}(Cry2)}$	0.5	2.236 [17]	7.813	3.484
4	PER protein peak to trough ratio $\frac{y_{\max}(\mathrm{PER})}{y_{\min}(\mathrm{PER})}$	5	> 20 [17]	large	large
5	CRY1 protein peak-trough ratio $\frac{y_{\max}(\mathrm{CRY1})}{y_{\min}(\mathrm{CRY1})}$	3	3.247 [17]	6.385	1.847
6	CRY2 protein peak-trough ratio $\frac{y_{\max}(\mathrm{CRY2})}{y_{\min}(\mathrm{CRY2})}$	3	1.975 [17]	8.094	2.347
7	Fraction PER of total protein $\frac{y_{\max}(\mathrm{PER})}{y_{\max}(\mathrm{PER}) + y_{\max}(\mathrm{CRY1}) + y_{\max}(\mathrm{CRY2})}$	3	0.105 [17]	0.169	0.073

	Description	Weight	Desired [ref]	Degradation	Activation
8	Fraction CRY1 of total protein $\frac{y_{\max}(\text{CRY1})}{y_{\max}(\text{PER}) + y_{\max}(\text{CRY1}) + y_{\max}(\text{CRY2})}$	3	0.555 [17]	0.473	0.554
9	Fraction CRY2 of total protein $\frac{y_{\max}(\text{CRY2})}{y_{\max}(\text{PER}) + y_{\max}(\text{CRY1}) + y_{\max}(\text{CRY2})}$	3	0.341 [17]	0.358	0.373
10	*Cry1* siRNA period sensitivity $\frac{\partial T}{\partial v_{\text{deg},\mathbf{c1}}}$	5	< 0 [22]	< 0	> 0
11	*Cry2* siRNA period sensitivity $\frac{\partial T}{\partial v_{\text{deg},\mathbf{c2}}}$	5	> 0 [22]	> 0	< 0
12	*Cry1* knockout period $\frac{T(Cry1^{-/-})}{T(\text{WT})}$	5	$< 95\%$ [21]	92.1%	193.4%

	Description	Weight	Desired [ref]	Degradation	Activation
13	*Cry2* knockout period $\frac{T(Cry2^{-/-})}{T(\mathrm{WT})}$	5	$>115\%$ [21]	131.8%	99.4%
14	Fraction CRY1 entering nucleus $\frac{y_{\max}(\mathrm{CRY1}_n)}{y_{\max}(\mathrm{CRY1}+\mathrm{CRY1}_n)}$	1	0.40 [17]	0.195	0.059
15	Fraction CRY2 entering nucleus $\frac{y_{\max}(\mathrm{CRY2}_n)}{y_{\max}(\mathrm{CRY2}+\mathrm{CRY2}_n)}$	1	0.35 [17]	0.139	0.059
16	Time delay between nuclear repressors and mRNA $t_{\max}(\text{nuclear protein}) - t_{\max}(\mathrm{mRNA})$	3	75% [17]	0.850	0.848

	Description	Weight	Desired [ref]	Degradation	Activation
17	Time delay between mRNA and cytoplasmic protein $t_{\max}(\text{mRNA}) - t_{\max}(\text{protein})$	3	25% [17]	0.057	0.095
18	Time delay between cytoplasmic protein and nuclear protein $t_{\max}(\text{protein}) - t_{\max}(\text{nuclear protein})$	3	0% [17]	0.093	0.057

2.2.3 Model validation and dynamics

The model was validated by comparing the simulated dynamics to novel experimental measurements. First, the time course plots of the state variables display reasonable phases and amplitudes, and oscillate with a period of 23.7 hours (figure 2.4). The knockout periods are 21.7 hr (91.4% of WT) for *Cry1*$^{-/-}$ and 31.5 hr (133% of WT) for *Cry2*$^{-/-}$, indicating that the two feedback loops are indeed redundant with different free-running periods.

SiRNA knockdowns

SiRNA knockdowns were performed *in silico* by increasing the degradation rate for the corresponding mRNA (figure 2.5), as done previously in [1]. *Cry1* and *Cry2*

Table 2.2: Best-fit parameter values. Best-fit parameters for the two alternative models proposed for the core clock feedback loop, as determined by fitting the cost function entries in table 2.1.

	Parameter	Description	model 1	model 2
1	$v_{\text{txn},\mathbf{p}}$	*Per* Transcription rate	0.195	0.276
2	$v_{\text{txn},\mathbf{c1}}$	*Cry1* Transcription rate	0.131	0.062
3	$v_{\text{txn},\mathbf{c2}}$	*Cry1* Transcription rate	0.114	0.053
4	$k_{\text{txn},\mathbf{p}}$	*Per* Repression constant	0.425	0.425
5	$k_{\text{txn},\mathbf{c}}$	*Cry1/2* Repression constant	0.259	0.262
6	$v_{\text{deg},\mathbf{p}}$	*Per* Max degradation rate	0.326	0.472
7	$v_{\text{deg},\mathbf{c1}}$	*Cry1* Max degradation rate	0.676	0.322
8	$v_{\text{deg},\mathbf{c2}}$	*Cry2* Max degradation rate	0.608	0.290
9	$k_{\text{deg},\mathbf{p}}$	*Per* Degradation constant	0.011	0.024
10	$k_{\text{deg},\mathbf{c}}$	*Cry1/2* Degradation constant	1.149	0.809
11	$v_{\text{deg},\mathbf{P}}$	Max PERc degradation rate	2.970	2.970
12	$k_{\text{deg},\mathbf{P}}$	PERc degradation constant	0.034	0.034
13	$v_{\text{deg},\mathbf{C1}}$	Max CRY1c degradation rate	1.523	1.048
14	$v_{\text{deg},\mathbf{C2}}$	Max CRY2c degradation rate	1.686	1.134
15	$k_{\text{deg},\mathbf{C}}$	CRYc degradation constant	2.017	2.028
16	$v_{\text{deg},\mathbf{CP}}$	CRYn degradation rate	0.101	0.070
17	$m_{\mathbf{C2N}}$	CRY2n degradation multiplier	3.318	3.334
18	$k_{\text{deg},\mathbf{CP}}$	CRYn degradation constant	0.053	0.053
19	$v_{\text{a},\mathbf{CP}}$	CRYn association rate	0.041	0.028
20	$v_{\text{d},\mathbf{CP}}$	CRYn dissociation rate	0.002	0.001
21	$k_{\text{tln},\mathbf{p}}$	PER translation rate	3.000	1.000

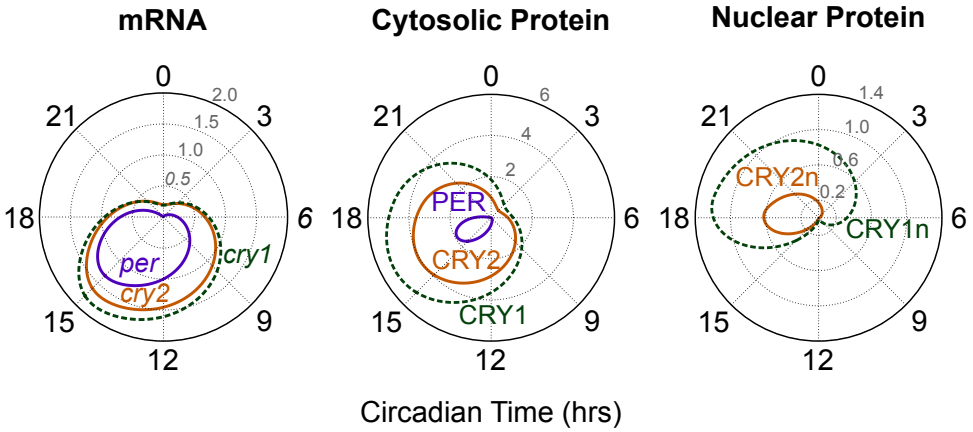


Figure 2.4: Time course concentration profiles of model 1. The three polar plots show the time-varying levels of each clock component for optimal parameter set. In these plots, the amplitude of each variable (always positive) is plotted against its phase, $0 \rightarrow 2\pi$. Since the limit cycle is periodic, this results in a closed curve. Rising mRNA levels caused by low repressor concentrations (CT12) result in accumulating cytosolic protein (left plot). Lower levels of PER prevent all the available CRY from entering the nucleus (middle plot). High levels of nuclear repressors halt transcription until both CRYs have degraded (right plot). Figure reprinted from [49].

knockdowns in wild type conditions show close agreement to experiment [22], and demonstrate that altering the nuclear CRY1/CRY2 ratio is effective in changing the period of oscillation.

CRY2 cytosolic stabilization

The *Dyrk1a* knockdown in [55] serves as another demonstration of the change in period as a response to a change in nuclear CRY1/CRY2 ratio. To demonstrate the effect in silico, the degradation rate of cytosolic CRY2 was decreased. Matching experimental evidence, the cytosolic CRY2 levels rose, with a corresponding change in the nuclear CRY ratio and decrease in period.

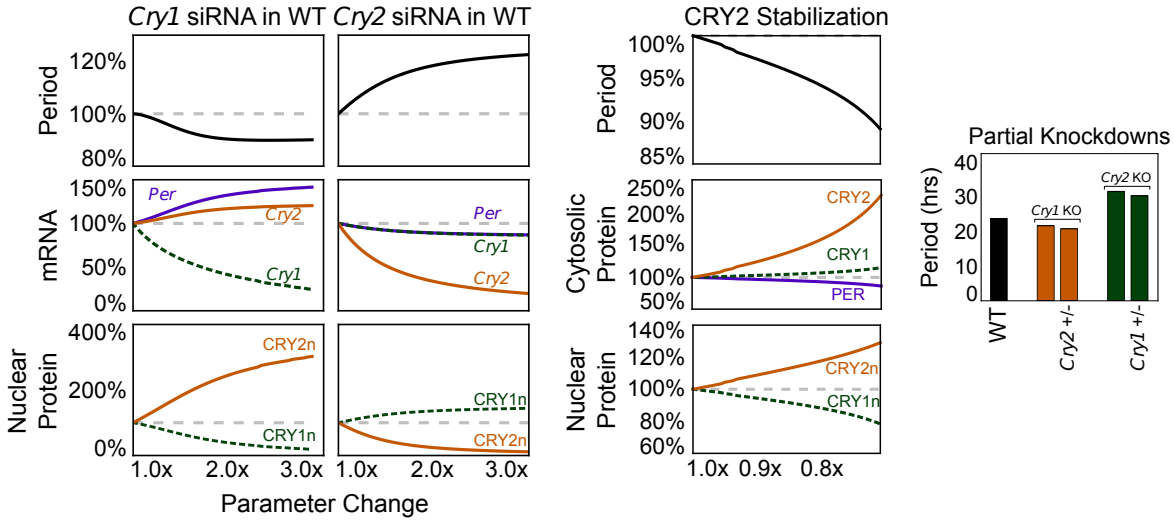


Figure 2.5: Validation of model 1 against experimental results. When comparing against existing experimental results, the model shows the correct response for *Cry1* siRNA, *Cry2* siRNA, and CRY_c stabilization (left, middle), through an adjustment in the nuclear CRY1/CRY2 ratio. The model also correctly captures the single/double knockout phenotype of [21], (right)

Single cry perturbations

The $Cry1^{+/-}$ $Cry2^{-/-}$ and $Cry1^{-/-}$ $Cry2^{+/-}$ single/double knockout perturbations from [21] were approximated with an appropriate knockout and siRNA knockdown, and show the correct period shortening. With one *Cry* knocked out, the levels of cytoplasmic PER are no longer stoichiometrically limiting. In this case, less *Cry* leads to less nuclear complex, which is cleared faster.

2.2.4 Prediction of KL001 mechanism

Possible mechanisms by which KL001 might lengthen the circadian period were investigated using the completed model. Experimental evidence indicates that stabilization of CRY can either lengthen the period (*Fbxl3* knockdown) or shorten the period (*Dyrk1a* knockdown). Assuming equal effect on both isoforms of CRY, the

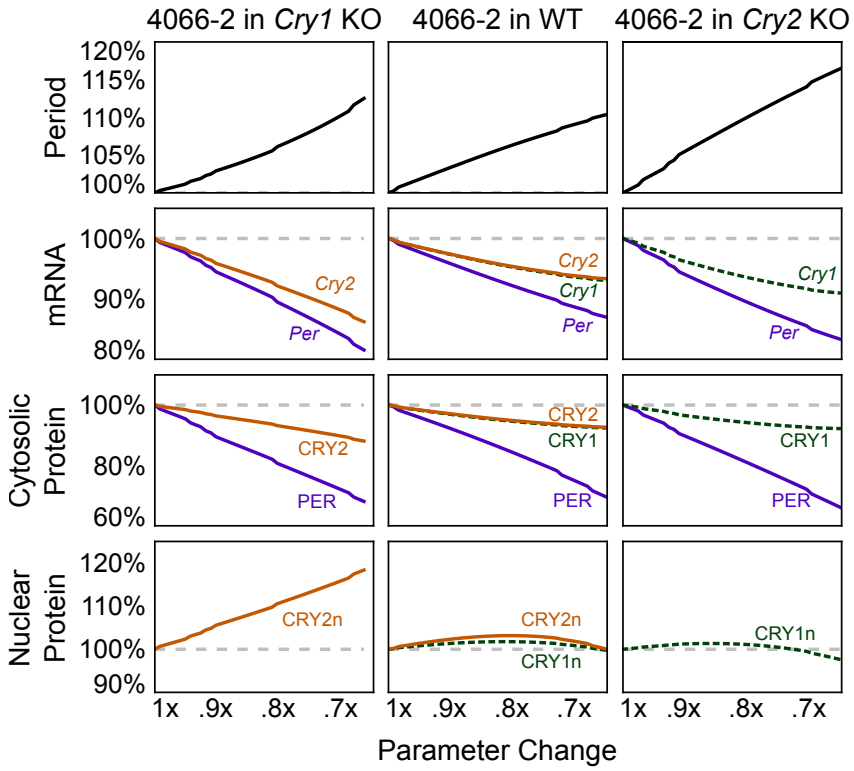


Figure 2.6: Prediction of the KL001 mechanism. *In silico* response to equal stabilization of both CRYs in the nucleus results in longer periods (middle column), matching observed results. The model correctly predicts that stabilization of individual CRYs in the knockout environments also causes period lengthening. Figure reprinted from [49].

model confirms that cytoplasmic stabilization results in period shortening, since excess CRY expedites the nuclear import of the PER:CRY repressive complex. However, stabilizing nuclear CRY (figure 2.2, last column) yields the appropriate period lengthening observed experimentally. We therefore predicted that KL001 acts primarily in the nucleus, potentially through the FBXL3 degradation pathway. Additionally, the model predicts that the compound's effect on $Cry1^{-/-}$ and $Cry2^{-/-}$ systems would be similar to its effect on the wild type clock, causing dose-dependent period lengthening (figure 2.6, left and right columns).

Experimental confirmation

Experimental tests of the model's predictions were performed. Nuclear CRY1 and CRY2 levels were up-regulated and almost sustained, respectively, while PER1 level was strongly down-regulated by the compound, supporting stabilization of nuclear CRY, the predicted mechanism. Additionally, continuous treatment with KL001 lengthened the period in both *Cry1* and *Cry2* knockout cells in a dose-dependent manner. Similarly, the compound caused period lengthening in both CRY1 and CRY2 knockdown U2OS cells. We further investigated the effect of KL001 in SCN explants, which show robust rhythms even in the absence of *Cry1*, due to intercellular coupling [61]. Both *Cry1* and *Cry2* knockout SCN exhibited dose-dependent period lengthening by the compound treatment. Thus, in a single *Cry* knockout, stabilization of either nuclear CRY1 and CRY2 causes period lengthening, confirming model predictions. Furthermore, the period-lengthening response of KL001 was abolished in cells treated with FBXL3 siRNA, further confirming that the period effect of KL001 is mediated by inhibiting FBXL3-CRY interactions (figure 2.7). This prediction was subsequently verified via structural biology, in which KL001 was shown to interfere directly with the binding of FBXL3 and CRY1 [62].

2.3 Discussion

2.3.1 Insights into circadian network design

Mathematical modeling has revealed how the balancing of two redundant feedback loops can provide fine control over the oscillatory period. A contour plot of period vs CRY1 or CRY2 abundance is shown in figure 2.8, which shows that the model's period is largely insensitive to the total amount of CRY (45° line), but highly de-

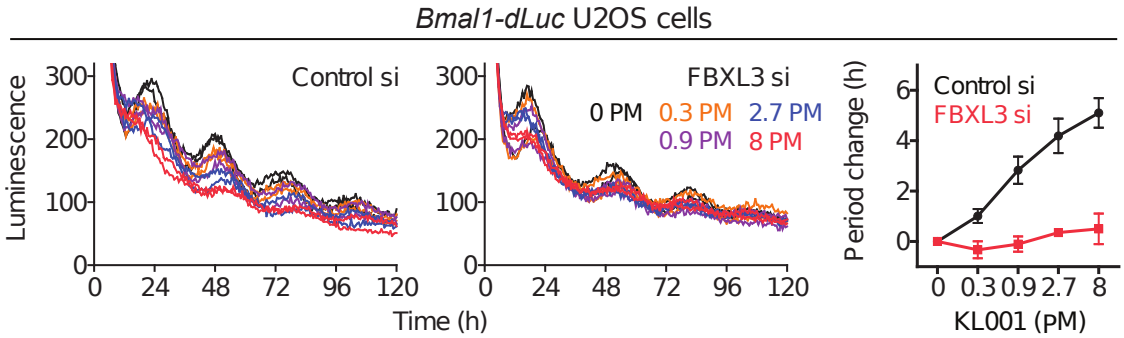


Figure 2.7: FBXL3 knockdown interferes with period-lengthening effect of KL001. The dose-dependent period increase of KL001 (right, black line) is abolished in cells treated with FBXL3 siRNA, indicating the mechanism of KL001 acts through FBXL3 and confirming model predictions. Figure reprinted from [49].

pendent on the ratio of the two isoforms. Sensitivity analysis of our mathematical model (figure 2.9) reveals that subtle changes in most of the involved rates can have an effect on the clock's free running period, which could be caused by evolutionary noise. Since many studies have linked entrainment to non-natural periods with long-term health problems, a mechanism to align the clock's natural period to that of the environment would be advantageous. It is therefore possible that the period control afforded by CRY1/CRY2 balancing is a deliberately conserved design principle of the circadian clock, which confers period robustness against random mutations of clock components. Indeed, the presence of independent control of the cytosolic CRY2 levels suggests that the biological clock can control its period through simple post-translational modifications [55].

2.3.2 Summary

Due to the high dimensionality of the model equations and sparsity of the experimental data, it is likely that more than one set of parameters would fit the cost function defined in table 2.1. While the fact that *any* parameter set for the given

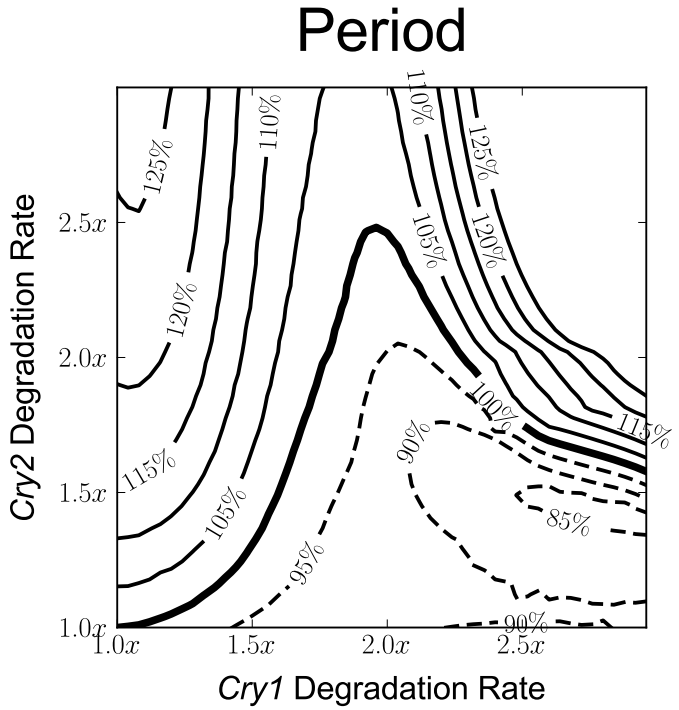


Figure 2.8: Simultaneous knockdown of *Cry1* and *Cry2*. Period contours (% change) resulting from simultaneous *in silico* siRNA knockdowns of *Cry1* (x-axis) and *Cry2* (y-axis). Since moving along the $y = x$ line remains close to the 100% period contour, the CRY1/CRY2 ratio largely determines the period, independent of total CRY, for perturbations up to twice the normal mRNA degradation rate.

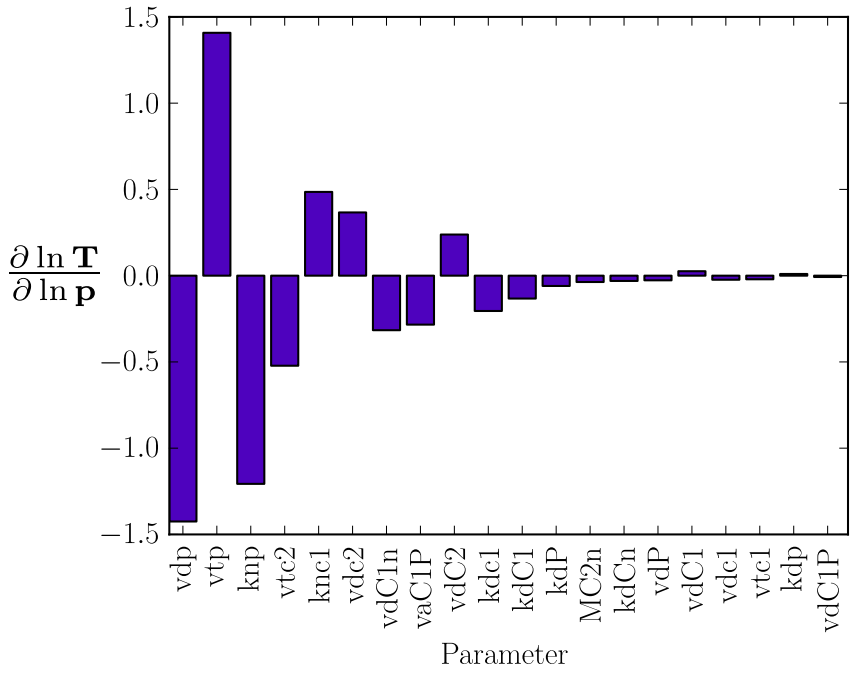


Figure 2.9: First order relative period sensitivities. Dimensionless period sensitivities with respect to the kinetic parameters in the degradation-based model. Perturbations to most clock rates result in noticeable changes to the free-running period. Many of the most sensitive parameters are transcriptional rates, which can be easily changed through mutation of DNA promoter regions.

model equations reproduces experimental results is enough to confirm that our hypothesized mechanism is mathematically consistent, we must be careful in placing too much confidence in predictions based on the particular parameter set generated by the optimization algorithm. While subsequent experimental results confirm many of our model predictions, a more rigorous exploration of parameter space would allow us to determine which predictions are constrained by the available experimental data. However, incorporating error ranges via standard techniques such as bootstrap techniques are infeasible when using a genetic algorithm optimization strategy. In chapter 3, we describe techniques to improve computational efficiency to allow a more systematic evaluation of predictive confidence.

Chapter 3

Identifiability analysis for models of circadian rhythms[1]

3.1 Background

A cell's behavior is governed by the dynamic and selective expression of its genes, in which each protein's activity depends on a careful balance between transcription, translation, transport, and degradation rates. These rates, which change with environmental conditions and are often impossible to measure accurately *in vivo* or *in vitro*, determine the function of a regulatory pathway. While studying the roles of individual proteins can often provide some insight on how a particular function is achieved, this approach is limited in explaining complicated cellular phenomena at the scale of dozens to hundreds of interacting genes. With the aid of mathematical models, it is increasingly possible to create *in silico* realizations of genetic regulatory networks to examine their dynamic properties. For example, the previous chapter

[1] Portions of this chapter are published in P. C. St. John and F. J. Doyle III, "Estimating confidence intervals in predicted responses for oscillatory biological models.," *BMC Syst. Biol.*, vol. 7, p. 71, July 2013.

presented a mathematical model of gene regulation for circadian rhythms that used new data from small molecule modulators to gain further insight into clock dynamics.

Essential to understanding how genetic circuits operate is connecting how inputs (i.e., environmental changes, extracellular signals) are processed to give the appropriate outputs (protein expression, cellular response). Models of genetic regulatory networks, often sets of ordinary differential equations (ODEs), contain many unknown parameters that must be estimated from experimental data [63]. Derivatives of the model output with respect to changes in input, known as local sensitivities, are frequently validated experimentally or used to predict potential targets for pharmaceuticals [64]. In chapter 2, local sensitivities were used within the cost function to ensure the model matched previously known experimental results. However, since sensitivities can change drastically with respect to the particular parameter values chosen, the confidence associated with parameter and sensitivity values is an important consideration in model analysis and design. were used

3.1.1 Identifiability analysis

Practical identifiability analysis is concerned with calculating confidence intervals in parameter estimates resulting from uncertainty in experimental data [65]. Several techniques for such an analysis currently exist, and are commonly used in analyzing biological models [66, 67, 68]. In one method, the inverse of the Fisher information matrix is used to provide estimates of the variance in each parameter. However, since this method assumes a linearized model, the resulting symmetric normal distributions for each parameter do not accurately reflect the mapping of nonlinear models [69]. In the bootstrap method, distributions in parameter estimates are found through optimum fits to repeated physical or *in silico* measurements. While accurate

in finding the true nonlinear confidence intervals, this approach requires efficient and robust parameter estimation convergence.

Difficulties imposed by periodic systems

In periodic systems, such as the models examined in this thesis, the behavior (and existence) of limit cycle oscillations is a discontinuous function of the parameters. Optimal values are traditionally found through trial-and-error type approaches [2, 43] or genetic algorithm search strategies [44], both of which are not amenable to bootstrap methods. Additionally, since the solutions are oscillatory, additional care must be taken in the calculation of the first-order sensitivity values. In this chapter, we calculate the sensitivity of the oscillatory period to parameter perturbation, a biologically relevant quantity that is often measured experimentally [31]. Due to these complications, rigorous identifiability analyses of these models are typically not performed.

In this chapter, a bootstrap uncertainty analysis appropriate for oscillatory biological models is developed and applied to the model of circadian rhythms described in chapter 2 [49]. These networks serve as an excellent example of a functional genetic circuit, able to process subtle environmental cues while remaining robust to temperature variations and evolutionary disturbances. Accurate limit cycle models must capture not only the correct time-dependent dynamics, but also the correct input-output response. For circadian rhythms, high-throughput microarrays have provided high-resolution time-series data of gene expression levels [70]. Additionally, knockdown experiments using RNA interference technology (siRNA) and small molecule modulators have resulted in a wealth of data on the dynamic responses to changes in key rates [22, 23, 49, 71]. This data, together with qualitative knowledge of the underlying network structure, permits the use and verification of a suitable

uncertainty analysis.

Alternative parameter estimation methods

To enable a bootstrap approach, we employ an efficient parameter estimation routine optimized for limit cycle models. Motivated by the increasing availability of high-resolution time-series measurements, we use an approach similar to multiple shooting, in which a nonlinear and discontinuous parameter estimation problem is transformed into a high-dimensional yet local optimization and solved via nonlinear programming [72]. Since the desired shape of the limit cycle solution is known *a priori*, a relatively accurate initial guess for the parameters and trajectories can be found. By using multiple sets of *in silico* data of varying quality, we illustrate how error in experimental results is propagated to uncertainty in parameter sensitivity. Lower quality data - with either higher error or fewer sampling points - result in wider distributions of limit cycles and less identifiable responses. These results can be used in *a priori* experimental design, finding the minimum sampling points needed for an estimated experimental error to enable accurate modeling. Additionally, we show using literature data how this method can be used to discriminate between candidate model structures, revealing which one yields the highest predictive confidence.

3.2 Methods

3.2.1 Collocation methods

The estimation of the unknown kinetic parameters is accomplished via nonlinear programming [72]. In this method, we divide the limit cycle trajectory $\mathbf{x}(t, \mathbf{p})$ into $\mathcal{N}$ finite elements of length h, and approximate each with a $\mathcal{K}$ degree Lagrange

interpolating polynomial, $\mathbf{x}_i^{\mathcal{K}}(t)$, using an internal time $\tau \in [0,1]$. For finite element i:

$$\begin{aligned} t &= h(i+\tau) \\ \ell_j(\tau) &= \prod_{k=0,\neq j}^{\mathcal{K}} \frac{\tau - \tau_k}{\tau_j - \tau_k} \\ \mathbf{x}_i^{\mathcal{K}}(\tau) &= \sum_{j=0}^{\mathcal{K}} \ell_j(\tau)\ \mathbf{x}_{ij}. \end{aligned} \tag{3.1}$$

We also ensure that the interpolating polynomial matches system dynamics at each collocation point, τ_k, by setting

$$\begin{aligned} \sum_{j=0}^{\mathcal{K}} \mathbf{x}_{ij} \frac{d\ell_j(\tau_k)}{d\tau} &= h\mathbf{f}(\mathbf{x}_{ij}, \mathbf{p}) \\ \text{for} \quad k &= 1, \ldots, \mathcal{K}. \end{aligned} \tag{3.2}$$

Additionally, the interpolating polynomials for each finite element must form a continuous function, so the following continuity constraints are imposed:

$$\begin{aligned} \mathbf{x}_{i+1,0} &= \sum_{j=0}^{\mathcal{K}} \ell_j(1)\ \mathbf{x}_{i,j} \\ \text{for} \quad i &= 1, \ldots, \mathcal{N}-1. \end{aligned} \tag{3.3}$$

Periodic conditions are imposed by setting the beginning of the first element equal to the end of the final element:

$$\mathbf{x}_{0,0} = \sum_{j=0}^{\mathcal{K}} \ell_j(1)\ \mathbf{x}_{\mathcal{N},j} \tag{3.4}$$

The τ_k values are chosen for optimal accuracy, here we use Gauss-Radau roots so

that the resulting method has stiff decay [72]. With $\mathcal{K} = 5$:

$$\tau = \{0.000, 0.057, 0.277, 0.584, 0.860, 1.000\}$$

The interpolating polynomials can now be compared to the experimental data. For each measured value, $\hat{\mathbf{x}}(t_n)$, the corresponding simulated values $\mathbf{x}^{\mathcal{K}}(t_n)$ can be interpolated from $\mathbf{x}_{ij}$:

$$\begin{aligned} \mathbf{x}^{\mathcal{K}}(t_n) &= \sum_{j=0}^{\mathcal{K}} \ell_j(\tau_n)\ \mathbf{x}_{ij} \\ \text{for} \quad n &= 1, \ldots, \mathcal{M} \end{aligned} \tag{3.5}$$

where i and τ_n are selected for the appropriate finite element and sampling time. The objective function $\Phi(\mathbf{x}, \mathbf{p})$ is thus:

$$\Phi(\mathbf{x}, \mathbf{p}) = \sum_{n}^{\mathcal{M}} \left(\frac{\mathbf{x}^{\mathcal{K}}(t_n) - \hat{\mathbf{x}}(t_n)}{\mathbf{s}_n} \right)^2 \tag{3.6}$$

where $\mathbf{s}_n$ is the measurement error associated with measurement n. Since $\mathbf{x}$ and $\mathbf{s}$ are vectors, the division in equation (3.6) must be performed element-wise. This cost function was taken from a similar multiple-shooting approach to parameter estimation [73].

Since the cost function (equation (3.6)) and equality constraints (equation (3.2), equation (3.3), and equation (3.4)) now satisfy continuity and differentiability requirements [74], parameter estimation can now be accomplished via constrained nonlinear programming (NLP) instead of a global search strategy. The solution is subject to variable bounds:

$$\begin{aligned} \mathbf{x}_{LB} \leq \mathbf{x} \leq \mathbf{x}_{UB} \\ \mathbf{p}_{LB} \leq \mathbf{p} \leq \mathbf{p}_{UB} \end{aligned} \tag{3.7}$$

The numerical implementation is accomplished using IPOPT [75], using the `MA57`

[76] linear solver. The CasADi computer algebra package [77] was used to provide an interface to the IPOPT numerical libraries and supply derivatives to the cost and equality function calls through automatic differentiation.

3.2.2 Generating initial values

Solution of the NLP described in section 3.2.1 requires a suitable initial guess for the optimal state profiles, $\mathbf{x}^\star$, and kinetic parameters, $\mathbf{p}^\star$. To find approximate values for these variables, a smoothed periodic B-spline, $\tilde{\mathbf{x}}$, is found using experimental data for each state variable using SciPy's interpolate module [78]. Initial values for x_{ij} are obtained by evaluating this spline at each τ_k for each finite element.

$$\begin{aligned} & \mathbf{x}^\star_{ij} \approx \tilde{\mathbf{x}}(t_{ij}) \\ \text{where} \quad & t_{ij} = h(i - 1 + \tau_j) \\ \text{for} \quad & i = \{1, \ldots, \mathcal{N}\},\ j = \{1, \ldots, \mathcal{K}\} \end{aligned} \tag{3.8}$$

Since

$$\frac{d\tilde{\mathbf{x}}}{dt} \approx \mathbf{f}(\tilde{\mathbf{x}}, \mathbf{p}), \tag{3.9}$$

approximate values for $\mathbf{p}$ can be obtained by solving the simpler unconstrained NLP,

$$\min_{\mathbf{p}} \sum_{i}^{\mathcal{N}} \sum_{j}^{\mathcal{K}} \left(\frac{d\tilde{\mathbf{x}}(t_{ij})}{dt} - \mathbf{f}(\tilde{\mathbf{x}}(t_{ij}), \mathbf{p}) \right)^2 \tag{3.10}$$

in which t_{ij} is the same as in equation (3.8) and the bounds on p are the same as in equation (3.7).

3.2.3 Numerical calculation of sensitivity coefficients

After determining an optimal parameter set for the given experimental data, relevant first order sensitivity coefficients for oscillatory models are found using the procedure from [31], summarized here. First, initial conditions and oscillatory period are verified by solving the boundary value problem (BVP):

$$\min_{\mathbf{x}(0),T} \begin{pmatrix} \mathbf{x}(T) - \mathbf{x}(0) \\ \dot{\mathbf{x}}_0(0) \end{pmatrix} \tag{3.11}$$

where $\dot{\mathbf{x}}_0(0)$ denotes the time-derivative of the first state variable, evaluated at $t = 0$. This BVP is solved using Newton's method, employing the SUNDIALS packages CVODES for ODE integration and KINSOL for the Newton iterations [60]. Time-dependent parametric sensitivities (equation (1.4)), are obtained by using the staggered-direct method from the CVODES integrator.

3.2.4 Generation of data for bootstrap methods

For each run, two thousand simulated measurements, $\hat{x}_i(t_j)$, were generated from the true data, $\tilde{x}_i(t_j)$, using a normal distribution with $\mu = \tilde{x}_i(t_j)$ and $\sigma_{ij} = \xi\ \tilde{x}_i(t_j) + \eta\ \max_j \tilde{x}_i(t_j)$, in which ξ is the relative and η is the absolute error. Each simulated data set was then used to find a unique optimum parameter set, $\mathbf{p}^\star$. Data sets that failed to converge, or reached a steady state solution (in which periodic sensitivities are undefined), were discarded from further analysis.

For the *in silico* data of varying quality used in figures 3.2-3.4, we used the known limit cycle $\mathbf{x}(t)$ to generate data points $\hat{x}_i(t_j)$ at each of $\mathcal{M}$ sampling points. The effect of increasing error and decreasing number sampling points were tested inde-

pendently:

$$\xi = \{.01, .05, .10, .20, .30\}; \quad \mathcal{M} = 20$$
$$\mathcal{M} = \{30, 20, 15, 10, 5\}; \qquad \xi = 0.15$$

Since standard deviations in the data distributions were also used as optimization weights, a small amount of absolute error ($\eta = 0.001$) was added to ensure errors in small values did not dominate the cost function.

3.2.5 Calculation times

Each parameter estimation took approximately 4 seconds on a 2.53GHz processor, with the subsequent limit cycle solution integration and sensitivity calculation taking approximately 0.5 seconds. Due to the parallel nature of the 2000 trials, computation times were alleviated by distributing the tasks onto a cluster of 160 compute nodes.

3.3 Results

Experimental values for the kinetic parameters of a limit cycle system, $\mathbf{p}$, are rarely available. Given time-series experimental measurements $\hat{x}_i(t_j)$ for each state variable in a limit cycle system, we find optimal parameters $\mathbf{p}^\star$ such that the error between the experimental measurements and the simulated limit cycle is minimized [73]:,

$$\mathbf{p}^\star := \arg\min_{\mathbf{p}} \sum_{i}^{states} \sum_{j}^{data} \frac{\left(\hat{x}_i(t_j) - x_i(t_j, \mathbf{p})\right)^2}{\sigma_{ij}^2}. \tag{3.12}$$

Here σ_{ij} is the standard deviation associated with the measured mean of state i at time j. Using the data points $\hat{x}_i(t_j)$ to generate a suitable initial guess, parameter

estimation may proceed via a nonlinear programming approach as described in section 3.2. In this chapter, we assume that all states are measured to demonstrate how initial guesses can be generated directly from the input data. However, for systems with unmeasured states, initial guesses for the trajectory and parameter values can be provided by another approach, such as a global optimization routine, as described in section 1.2.6. A bootstrap method was implemented by repeatedly sampling input data distributions to calculate a population of optimal parameter fits.

After finding optimal parameter fits, we used the models to predict how perturbations change systems dynamics by performing a first order sensitivity analysis. Since adjustments to periodic systems in response to inputs are often manifested through temporary changes in oscillatory period, relative period sensitivities, equation (1.7), were calculated due to their independence of parameter magnitude [30, 31, 79]. Relative period sensitivities were integrated into the bootstrap method by calculating appropriate sensitivities for each estimated parameter set.

Of particular importance in determining the reliability of a model prediction is whether an output response maintains a consistent direction despite noise in measurement data. We therefore define a sensitivity value to be practically identifiable for given input data if 95% of the distribution maintains a consistent sign, similar to definitions for parameter identifiability used in previous studies [69, 80]. An overview of the method is shown in figure 3.1.

3.3.1 Effect of data quality on predictive confidence

We first analyze the degree to which uncertainty in input data is propagated to uncertainty in output predictions. To achieve this, we generate *in silico* data from a previously published model of circadian rhythms (model 1), using relative error ξ to generate normally distributed data ($\sigma_{ij} = \xi\ \hat{x}_i(t_j)$) at each of $\mathcal{M}$ sampling points. As

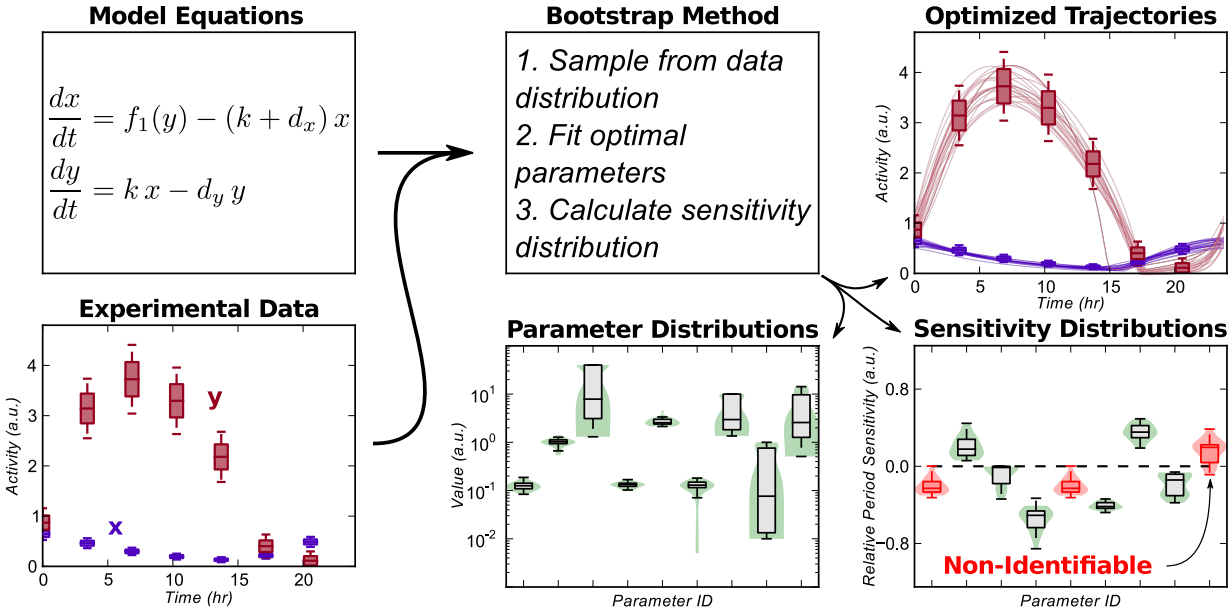


Figure 3.1: Parameter estimation and bootstrap methods flowchart. The demonstrated method calculates confidence intervals in the sensitivity of limit cycle models. An oscillatory model and experimental (or simulated) data are inputs to the bootstrap method. Unique data sets are then used to calculate optimum limit cycle trajectories. The resulting distribution in sensitivities highlight whether a particular response is identifiable (i.e., consistent across the majority of bootstrap trials). Figure reprinted from [81].

expected, solution trajectories drifted further from the nominal limit cycle for higher values of error, ξ, or lower sampling density, $\mathcal{M}$, (figure 3.2). However, the overall shape of the oscillatory profiles remained relatively similar, even for rather high ξ or low $\mathcal{M}$.

Figure 3.3 shows violin plots of the probability distribution for each parameter set and corresponding sensitivity evaluation for increasing ξ, while figure 3.4 shows similar plots for decreasing $\mathcal{M}$. Interestingly, there is little correlation between the identifiability of a parameter and its corresponding sensitivity value. For example, vdP, the maximum degradation rate of Per mRNA, shows a very tight clustering about its nominal parameter value, while the sensitivity of this parameter loses identifia-

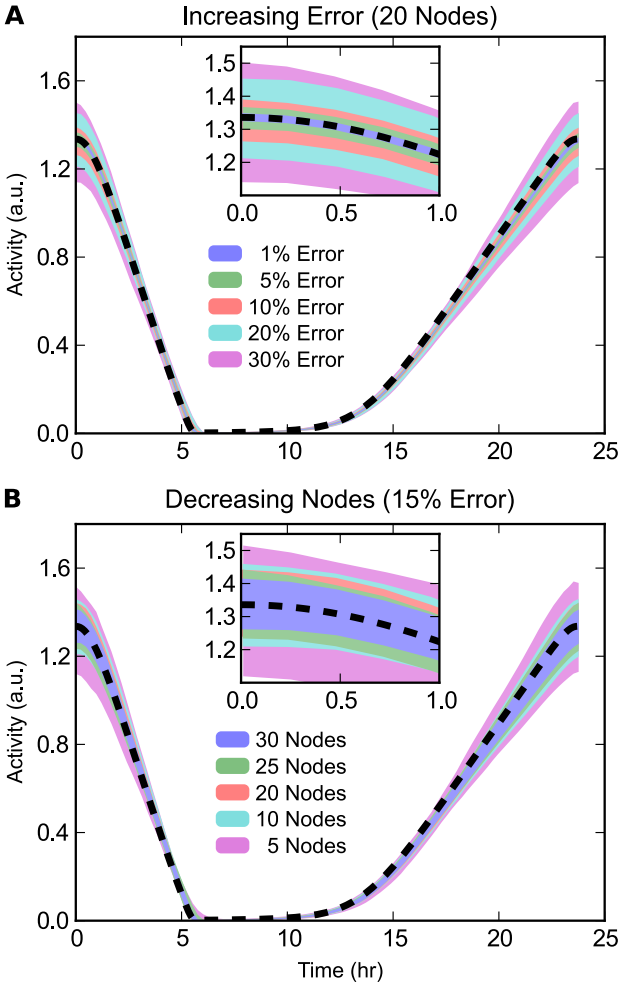


Figure 3.2: Time-course profiles of the state trajectories for *Per mRNA*. (**A**) Increasing relative error, ξ, with $\mathcal{M} = 20$. Possible state variable values are shown as shaded regions, obtained by filling between the 5^{th} and 95^{th} percentile for values at each time for 2000 independent parameter estimations. Increasing ξ results in larger deviations from the original model trajectory, shown as a dashed black line. (**B**) Decreasing number of measurement points, $\mathcal{M}$, each with $\xi = 0.15$. Higher $\mathcal{M}$ results in trajectories closer to the true trajectory. Figure reprinted from [81].

bility for even small values of ξ. Conversely, KdCn, the Michealis-Menten constant associated with the degradation of nuclear CRY, shows large variations in possible parameter values. However, the period sensitivity of KdCn, despite lying close to the x-axis, remains identifiable, indicating a robust prediction. These results reveal which model responses are constrained by the structure and dynamics of the limit cycle oscillations, and which are dependent on the particular parameterization chosen.

Sensitivities that are experimentally distinguishable from zero are the most important for validation. Calculating a typical experimental value for a relative period sensitivity helps to calibrate which sensitivities might be verified experimentally. Referring to a recent RNA interference screen, periods changes of approximately 1 hour (5%) can be reliably measured using luminescence recordings [22]. Assuming an increase in the corresponding mRNA degradation parameter value of 50%, this translates to a relative period sensitivity of 0.1. Thus, many of the identifiable values shown in figures 3.3-3.4 fall within the experimentally measurable range.

3.3.2 Application to literature data for model discrimination

We next apply the method to literature time-course data for core clock components [17]. When modeling a genetic regulatory network, many candidate model equations are often considered. We show that a bootstrap uncertainty analysis can also be useful in discriminating between potential model structures based on predictive confidence. Here two variations of the same model are fit: The first model (model 1) was originally optimized using a genetic algorithm approach, and thus contains a minimal number of parameters to reduce optimization complexity. The second model (model 3) considered contains independent parameters for each rate expression, increasing the number of parameters from 23 to 35.

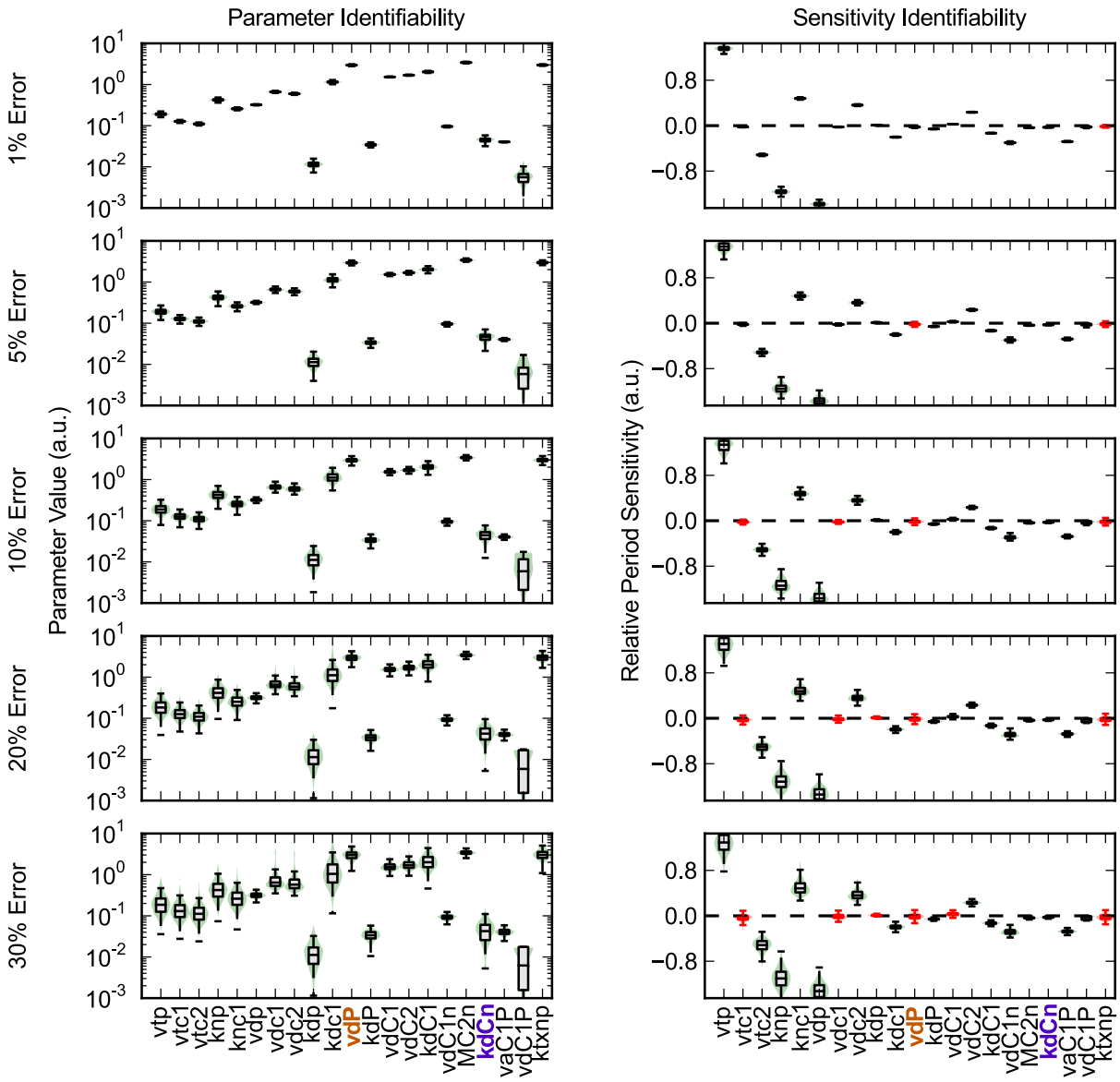


Figure 3.3: Parameter and sensitivity identifiability for increasing error. Increasing ξ results in a corresponding decrease in the confidence of the parameter and sensitivity estimates. Violin plots of the parameter values (left) and relative period sensitivities (right) show the distribution of values from each parameter estimation. In the plots, a box plot is superimposed above a kernel density plot to convey the distribution of values. The whiskers used extend to the most extreme data point within 1.5x the inner quartile range. Sensitivities in which the 5^{th} and 95^{th} percentile values span the x-axis are deemed non-identifiable (red), as the model's response direction can not be accurately estimated. Higher ξ also results in wider parameter distributions. Figure reprinted from [81].

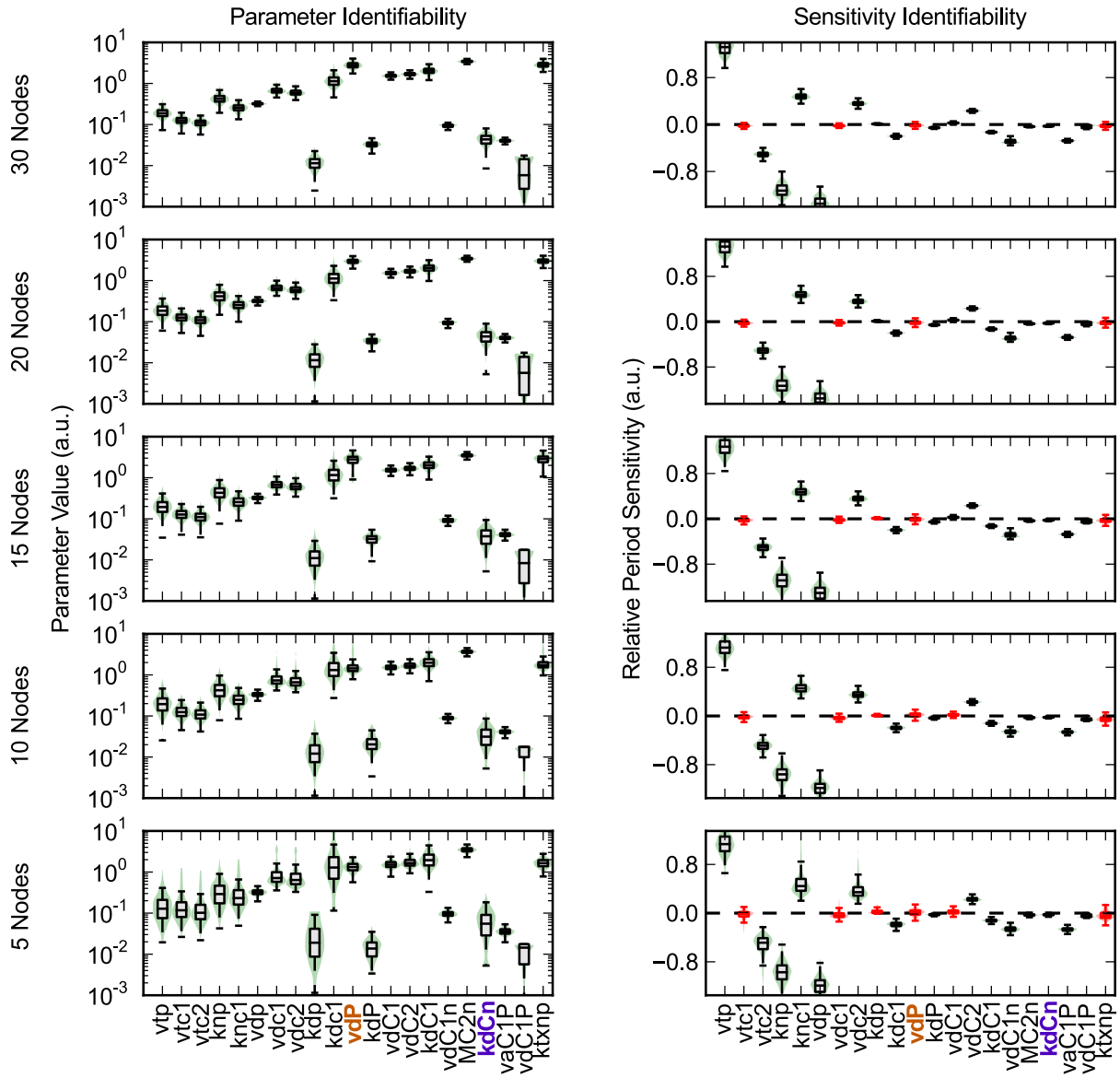


Figure 3.4: Effect of high-resolution sampling on identifiability. Lower values of $\mathcal{M}$ result in less constrained parameter and sensitivity values. Similar to figure 3.3, violin plots of the parameters (left) and sensitivities (right) show the distribution from each parameter estimation for decreasing $\mathcal{M}$. These results highlight the importance of high-resolution time sampling in generating sensitivity information for oscillatory models. Figure reprinted from [81].

Table 3.1: Parameter values for model 3. Parameters were fit to time-series data via nonlinear programming.

	Parameter	Description	Value
1	M_1	*Per*/CRY2 activity coefficient	3.632
2	Vm_1	*Per* transcription rate	9.957×10^{-1}
3	Ki_1	*Per*/CRY inhibition coefficient	1.054×10^{-1}
4	M_2	*Cry1*/CRY2 activity coefficient	1.000×10^{-3}
5	Vm_2	*Cry1* transcription rate	2.262×10^{-1}
6	Ki_2	*Cry1*/CRY inhibition coefficient	2.049×10^{-1}
7	M_3	*Cry2*/CRY2 activity coefficient	1.000×10^{1}
8	Vm_3	*Cry2* transcription rate	1.850×10^{-1}
9	Ki_3	*Cry2*/CRY inhibition coefficient	1.427×10^{-1}
10	k_1	*Per* degradation rate	2.920×10^{-1}
11	Km_1	*Per* degradation self-inhibition	6.609×10^{-1}
12	k_2	*Cry1* degradation rate	1.000×10^{1}
13	Km_2	*Cry1* degradation self-inhibition	1.823×10^{-2}
14	k_3	*Cry2* degradation rate	4.711×10^{-2}
15	Km_3	*Cry2* degradation self-inhibition	1.000×10^{1}
16	k_4	*Per* translation rate	1.132×10^{-1}
17	k_5	*Cry1* translation rate	3.409×10^{-1}
18	k_6	*Cry2* translation rate	1.961×10^{-1}
19	k_7	PER degradation rate	1.000×10^{1}
20	Km_4	PER degradation self-inhibition	1.911×10^{-3}
21	k_8	CRY1 degradation rate	1.000×10^{1}
22	Km_5	CRY1 degradation self-inhibition	2.077×10^{-2}
23	k_9	CRY2 degradation rate	1.000×10^{1}
24	Km_6	CRY2 degradation self-inhibition	1.180×10^{-2}
25	k_{10}	C1N association rate	5.022×10^{-1}
26	k_{11}	C2N association rate	1.035
27	k_{12}	C1N dissociation rate	1.000×10^{-3}
28	k_{13}	C2N dissociation rate	2.070×10^{-1}
29	M_4	CRY1n/CRY2n activity coefficient	1.000×10^{1}
30	k_{14}	CRY1N degradation rate	1.000×10^{1}
31	Km_7	CRY1n degradation inhibition	5.206×10^{-3}
32	M_5	CRY2n/CRY2n activity coefficient	1.000×10^{1}
33	k_{15}	CRY2n degradation rate	1.000×10^{1}
34	Km_8	CRY2n degradation inhibition	1.829×10^{-2}

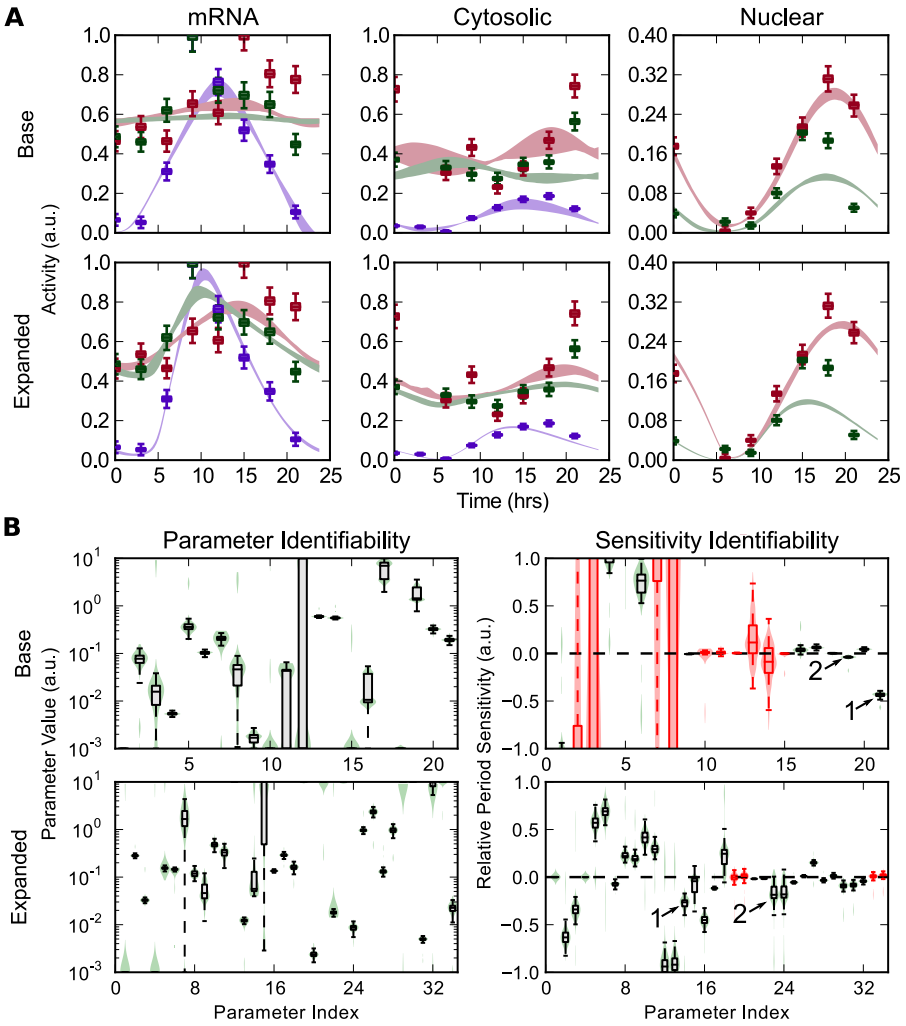


Figure 3.5: Identifiability comparison of two model structures. (**A**) Bootstrap parameter estimations on two model structures using literature time-series data with estimated errors (box plots). Resulting regions of model trajectories are shaded between the 5th and 95th percentile. *Per* species are shown in purple, *Cry1* in red, and *Cry2* in green. While both models were able to approximately reproduce the same dynamic response, the expanded model was better able to capture differences between the *Cry1* and *Cry2* profiles. (**B**) Parameter and sensitivity identifiability for the base and expanded models. Violin plots show the parameter and sensitivity distributions, with unidentifiable sensitivities (90% confidence level) highlighted in red. Despite containing more parameters, the expanded model shows better parameter identifiability and higher confidence in its predicted sensitivities. The PER translation rate (1) and PER-CRY association rate (2) sensitivities are consistent across model equations and are highlighted. Figure reprinted from [81].

Model 3: An expanded parameterization of model 1. These equations have similar reaction stoichiometry to those in model 1, but with more parametric degrees of freedom. This model showed better time-series performance than the more constrained model when fit to time-series data.

$$\frac{d\mathbf{p}}{dt} = \frac{Vm_1}{1 + Vm_1 \left(\frac{\mathbf{C1N}+\mathbf{C2N}\, M_1}{Ki_1}\right)^3} - \frac{k_1\ \mathbf{p}}{1 + \frac{\mathbf{p}}{Km_1}}$$

$$\frac{d\mathbf{c1}}{dt} = \frac{Vm_2}{1 + Vm_2 \left(\frac{\mathbf{C1N}+\mathbf{C2N}\, M_2}{Ki_2}\right)^3} - \frac{k_2\ \mathbf{c1}}{1 + \frac{\mathbf{c1}}{Km_2}}$$

$$\frac{d\mathbf{c2}}{dt} = \frac{Vm_3}{1 + Vm_3 \left(\frac{\mathbf{C1N}+\mathbf{C2N}\, M_3}{Ki_3}\right)^3} - \frac{k_3\ \mathbf{c2}}{1 + \frac{\mathbf{c2}}{Km_3}}$$

$$\frac{d\mathbf{P}}{dt} = k_4\ \mathbf{p} + k_{12}\ \mathbf{C1N} + k_{13}\ \mathbf{C2N} - \frac{k_7\ \mathbf{P}}{1 + \frac{\mathbf{P}}{Km_4}} - k_{10}\ \mathbf{P}\ \mathbf{C1} - k_{11}\ \mathbf{P}\ \mathbf{C2}$$

$$\frac{d\mathbf{C1}}{dt} = k_5\ \mathbf{c1} + k_{12}\ \mathbf{C1N} - \frac{k_8\ \mathbf{C1}}{1 + \frac{\mathbf{C1}}{Km_5}} - k_{10}\ \mathbf{P}\ \mathbf{C1}$$

$$\frac{d\mathbf{C2}}{dt} = k_6\ \mathbf{c2} + k_{13}\ \mathbf{C2N} - \frac{k_9\ \mathbf{C2}}{1 + \frac{\mathbf{C2}}{Km_6}} - k_{11}\ \mathbf{P}\ \mathbf{C2}$$

$$\frac{d\mathbf{C1N}}{dt} = k_{10}\ \mathbf{P}\ \mathbf{C1} - k_{12}\ \mathbf{C1N} - \frac{k_{14}\ \mathbf{C1N}}{1 + \frac{\mathbf{C1N}+\mathbf{C2N}\, M_4}{Km_7}}$$

$$\frac{d\mathbf{C2N}}{dt} = k_{11}\ \mathbf{P}\ \mathbf{C2} - k_{13}\ \mathbf{C2N} - \frac{k_{15}\ \mathbf{C2N}}{1 + \frac{\mathbf{C1N}+\mathbf{C2N}\, M_5}{Km_8}}$$

The literature data used consisted of 7-8 concentration time points across a 24 hour period. Confidence intervals in the data were not available, so an optimistic 3% relative and 0.5% absolute error was assumed for each data point ($\sigma_{ij} = 0.03\ \hat{x}_i(t_j) + 0.005\ \max(\hat{x}_i)$). Figure 3.5A shows the resulting time-series profiles for bootstrap estimations of each model. While additional kinetic parameters are typically thought to lower the predictive confidence of a model (the 'curse of dimensionality'), the expanded model is able to better capture the oscillatory profiles with lower variability between solutions. Parameter and sensitivity distributions, figure

3.5B, similarly show how the expanded model parameterization is able to generate more confident predictions in model response. Since the resulting sensitivity identifiability for both models was relatively poor, we highlight sensitivities which pass a 90% confidence level threshold. These results thus indicate higher-resolution data on circadian components would help in conferring confidence to model predictions.

Two sensitivities, the PER translation rate and the PER-CRY association rate, had high confidence and consistent direction in both the base and expanded parameterization - suggesting that the predicted values are robust to slight changes in both parameter value and model structure (figure 3.5B). Since a biological system can be modeled using many different combinations of kinetic assumptions, such a technique will likely prove useful in finding consistent predictions which are robust to slight differences in model equations.

Figure 3.6 compares optimal fits for both the base and expanded models to the originally published parameter set [49]. Since the original cost function was concerned mostly with optimizing stoichiometric and knockout data, the refitted models are able to much more accurately represent the time-series dynamics.

3.4 Discussion

Increasingly, mathematical models are being used to study biological systems where traditional experiments would prove infeasible. For example, in the search for drug targets, thousands of possible combinatorial perturbations can be quickly scanned for therapeutic effects using *in silico* modeling. This is especially useful in oscillatory systems with long periods, such as circadian rhythms, where a perturbed *in vitro* or *in vivo* system must be measured for multiple days before changes can be reliably determined.

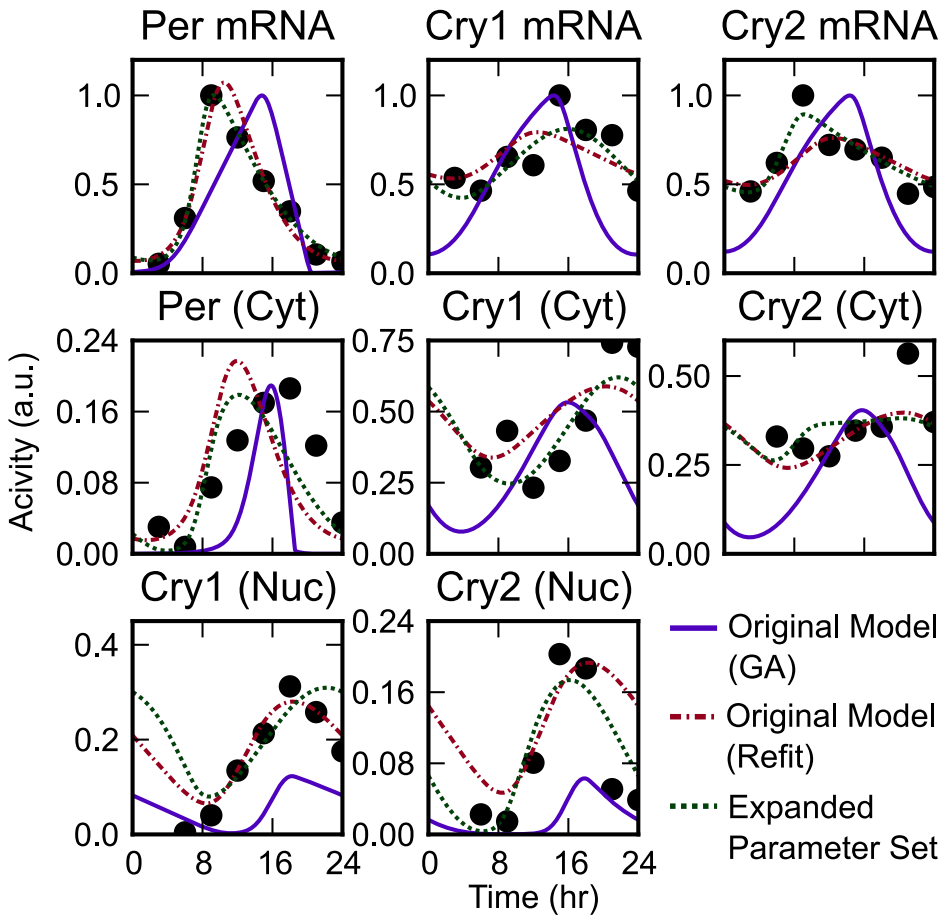


Figure 3.6: Time-series dynamics of fitted models. Model trajectories for each of the considered models. The original model (purple) shows the time series dynamics for the previously published parameter set, used in Figures 1-4. In these plots, the original model's period and amplitudes were rescaled to best match the experimental data, shown in black (without changing the dynamic profiles). The refitted model (red, dashed) was generated by optimizing the dynamic profile to time series data using the parameter estimation routine described in section 3.2.1. The expanded model (green, dashed) consists of a similar network structure, but with independent kinetic parameters for each rate expression (model 3). These plots show that parameter optimization through nonlinear programming is able to more accurately fit gene and protein expression profiles. Figure reprinted from [81].

However, since errors in model responses can arise from either incorrect structure or measurement noise, our confidence in *in silico* predictions is limited. Here a bootstrap approach suitable for periodic systems has been developed and extended to include uncertainty in predicted responses. With this method, errors due to local parameter effects can be identified, even in models with complicated dynamics. Furthermore, by considering multiple variations in model assumptions, we have demonstrated that more trustworthy model predictions can be found.

Since this method takes advantage of time-series data to generate a strong initial guess for an otherwise difficult parameter estimation, it requires high-resolution data on the concentrations of all species in the model. In many biological systems, such data is only available for the activity levels of certain well-studied species. However, the continued development of high-throughput genomic and proteomic techniques promise to deliver time-series data for a much larger network of components. With expanding datasets, these methods will likely prove useful for the quantitative evaluation of uncertainty in larger biological models.

In the following chapter, we apply this method to several existing circadian models in order find conserved predictions across different model kinetics and parameterizations. Such an approach allows us to further differentiate between the actions of two seemingly similar small molecule perturbations.

Chapter 4

Spatiotemporal separation of PER and CRY posttranslational regulation[1]

In the previous chapter, computationally efficient techniques for preforming a bootstrap analysis on circadian models was described. In this chapter, we apply this technique to clarify the roles of two posttranslational regulators in the core circadian feedback loop. Since direct experimental evidence of the systems-level effect of these regulators is difficult if not impossible to obtain, we must rely on computational modeling to provide insight into the dynamic consequences of modulating posttranslational activity. For this reason, bootstrap methods are invaluable in generating confident *in silico* results.

[1]Portions of this chapter are published in P. C. St. John, T. Hirota, S. A. Kay, and F. J. Doyle III, "Spatiotemporal separation of PER and CRY posttranslational regulation in the mammalian circadian clock.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, pp. 2040–5, Feb. 2014.

4.1 Background

Since circadian and metabolic regulators are tightly integrated, circadian disruptions often manifest in metabolic disease [82]. Recent efforts have therefore sought to gain a mechanistic understanding of these pathways, such that the metabolic burdens imposed by a 24-hour society might be mitigated. Post-translational regulators, which play key roles in connecting circadian and metabolic processes, serve as likely targets for future therapeutics – demonstrated by the wealth of available circadian-active small molecules [48, 83].

4.1.1 Nuclear entry of PER-CRY complex

Experimental evidence on PER/CRY nuclear entry is seemingly contradictory. For nuclear localization of the PER and CRY proteins, experiments from *Per1*$^{-/-}$ *Per2*$^{-/-}$ mice demonstrate that PER proteins are required for timely nuclear accumulation of CRY [17]. While other studies have shown the possibility of nuclear entry of CRY without PER, these results are typically based on steady-state endpoint assays which do not consider the speed of CRY nuclear entry [52, 84, 85]. We therefore consider the formation of the PER-CRY heterodimer as a key step in nuclear entry, which is supported by the fact that to the best of our knowledge all circadian models that consider both PER and CRY employ this kinetic assumption [1, 2, 43, 44, 49].

4.1.2 Importance of posttranslational regulators

The stabilities of PER and CRY are tightly regulated: PER proteins are phosphorylated by the casein kinase we family of proteins (CKIδ/ϵ), prompting β-TrCP-mediated degradation [86] and nuclear import [87]. The degradation of CRY pro-

teins is separately regulated by the SCF^{FBXL3} ubiquitin ligase complex [88, 89, 90]. The activities of both CKI-PER and FBXL3-CRY may be further coupled to the cell's metabolic state through AMPK signalling [91]. These posttranslational regulatory mechanisms have a strong effect on period length: the gain-of-function mutant $CKI\epsilon^{tau}$ leading to hyperphosphorylation of PER [92] and small molecule CKI inhibitors, such as longdaysin [71], demonstrated that increasing or decreasing CKI-dependent PER phosphorylation shortens or lengthens the period, respectively. In contrast, genetic mutations of FBXL3 [89, 90] and KL001, a small molecule inhibitor of FBXL3-dependent CRY degradation [49], showed that increased CRY stability leads to longer periods. Since the scale and complexity of the circadian network complicates an intuitive understanding of these relationships, mathematical models have played important roles in understanding how these manipulations affect circadian period [49, 86, 92].

4.1.3 Functional differences between PER and CRY

Given that both CKI and FBXL3 pathways similarly regulate the stability of linked negative factors, it was thought that simultaneous perturbations to both pathways might lead to non-additive effects: *i.e.*, the slowest link would determine the period [93]. This relationship was expected since the stabilities of both nuclear PER and CRY were thought to be important in the determination of clock kinetics [49, 94]. However, both small molecule [49] and genetic experiments [93] have demonstrated the independent period effects of CKI and FBXL3 post-translational regulations, as the inhibition of one pathway does not diminish the sensitivity of the other. This observation may be explained by a recent clarification of the canonical clock feedback circuit, where dissociated CRY was revealed as the dominant repressor of CLOCK-BMAL1 mediated E box transcription [84]. This distinction helps differentiate between the

roles of the otherwise similar PER and CRY proteins, in which the main role of PER in transcriptional repression is likely regulating the timing of CRY's nuclear accumulation. Therefore, while previous mathematical models in which PER acts as a direct repressor have proposed mechanisms for CKI-dependent period lengthening [92, 95], they are likely not suitable for distinguishing between CKI-PER and FBXL3-CRY mediated period change.

In this chapter, we used human cells harboring clock gene reporters together with mathematical modeling to gain insight into the relationship between PER and CRY posttranslational regulation. Consequently, we provide a new mechanism by which CKI-dependent PER phosphorylation controls the circadian period separately from the FBXL3-CRY pathway. The resulting detailed understanding of PER and CRY regulation in the core feedback loop provides a framework on which to interpret metabolic and pharmacological control of circadian rhythms.

4.2 Methods

4.2.1 Analysis of luminescence profiles

Raw luminescence data was first separated into a moving baseline and oscillatory component using a Hodrick-Prescott filter with a smoothing parameter of 1600. Example trajectory decompositions are shown in figure 4.1. Amplitudes (as shown in figure 4.3) were determined by taking the standard deviation in the baseline-subtracted data. Periods were obtained by nonlinear curve fitting, in which a four parameter (initial amplitude, decay, period and phase) damped cosine curve was fit to the baseline-subtracted data. Periods were not shown if the relative amplitude (found by standard deviation) fell below 25%, since noise dominated the periodic

trajectory.

4.2.2 Cost function

Models were fit to a cost function of experimental results. *Per*, *Cry*, *Clock*, and *Bmal1* protein and mRNA levels were taken from [17], along with profiles of CRY nuclear localization. For the model from [1], additional activity profiles on *Rev-Erb* and *Ror* were obtained from CircaDB (`http://bioinf.itmat.upenn.edu/circa/`). To score a model trajectory, mRNA state variables were scaled independently to minimize the squared error between model and experiment, since model parameters could be adjusted to give mRNA profiles arbitrary amplitudes. For protein species, where stoichiometric interactions are important, a single scaling parameter was used for all species. Nuclear repressor species, in which only relative measurements were available, were scaled independently. Full model equations for the model from Hirota *et al.*, 2012 are shown in model 1, the model from Relógio *et al.* 2011 in model 4, and the model from Leloup & Goldbeter, 2003 in model 5.

4.2.3 Parameter estimation and bootstrap analysis

Bootstrap parameter estimations were performed as described in chapter 3 with data from [17] assumed to have a normally distributed 10% relative and 5% absolute error. Since not all states in the models were measured, initial guess values for the trajectory and parameter variables were generated by optimizing the parameter sets first with a genetic algorithm approach, described in [44]. To help ensure bootstrap trials remained in a similar stability region of parameter space (and protect against steady-state solutions), bootstrap parameters were bound between 50% and 150% of their initial value.

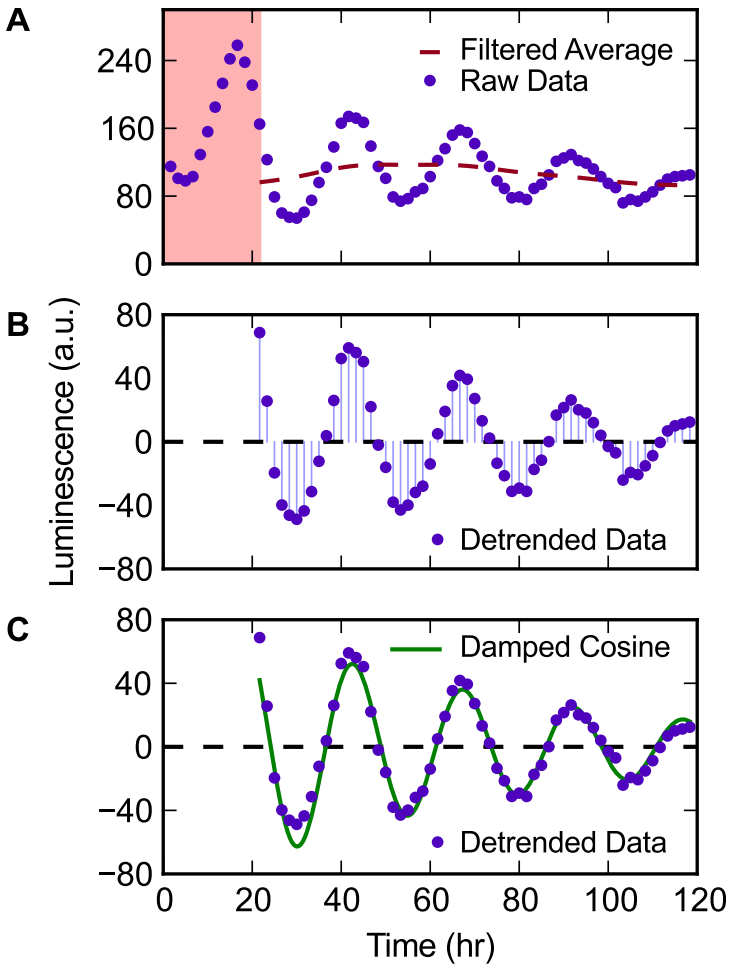


Figure 4.1: Analysis of circadian reporter luminescence data. (**A**) Raw luminescence data is first cropped by removing the initial transient region (first 12 points). The moving baseline is estimated using a Hodrick-Prescott filter with smoothing parameter 1600 (red dashed line). (**B**) Data is detrended by subtracting the moving baseline from the raw data. The detrended data is used to calculate the relative amplitude of the oscillations via standard deviation. (**C**) Periods are estimated by fitting a damped cosine curve (green solid line) to the detrended data. Figure reprinted from [96].

Model 4: Model from Relógio *et al.* 2011 [1].

$$\frac{d\,CLKBM1}{dt} = BM1n \cdot kfCLKBM1 - CLKBM1 \cdot dCLKBM1 - CLKBM1 \cdot kdCLKBM1$$

$$\frac{d\,reverb}{dt} = \frac{V3max\left(g \cdot \frac{CLKBM1}{kt3}^{v} + 1\right)}{\frac{CLKBM1}{kt3}^{v}\left(\frac{PnCn+PnpCn}{ki3}\right)^{w} + \frac{CLKBM1}{kt3}^{v} + 1} - dreverb \cdot reverb$$

$$\frac{d\,ror}{dt} = \frac{V4max\left(h \cdot \frac{CLKBM1}{kt4}^{p} + 1\right)}{\frac{CLKBM1}{kt4}^{p} \cdot (\frac{PnCn+PnpCn}{ki4})^{q} + \frac{CLKBM1}{kt4}^{p} + 1} - dror \cdot ror$$

$$\frac{d\,REVERBc}{dt} = -REVERBc \cdot dREVERBc - REVERBc \cdot kiREVERBc + kp3 \cdot reverb$$

$$\frac{d\,RORc}{dt} = -RORc \cdot dRORc - RORc \cdot kiRORc + kp4 \cdot ror$$

$$\frac{d\,REVERBn}{dt} = REVERBc \cdot kiREVERBc - REVERBn \cdot dREVERBn$$

$$\frac{d\,RORn}{dt} = RORc \cdot kiRORc - RORn \cdot dRORn$$

$$\frac{d\,bm1}{dt} = \frac{V5max\left(i \cdot \frac{RORn}{kt5}^{n} + 1\right)}{\frac{REVERBn}{ki5}^{m} + \frac{RORn}{kt5}^{n} + 1} - bm1 \cdot dbm1$$

$$\frac{d\,BM1c}{dt} = -BM1c \cdot dBM1c - BM1c \cdot kiBM1c + bm1 \cdot kp5$$

$$\frac{d\,BM1n}{dt} = BM1c \cdot kiBM1c - BM1n \cdot dBM1n - BM1n \cdot kfCLKBM1 + CLKBM1 \cdot kdCLKBM1$$

$$\frac{d\,per}{dt} = \frac{V1max\left(a \cdot \frac{CLKBM1}{kt1}^{b} + 1\right)}{\frac{CLKBM1}{kt1}^{b} \cdot (\frac{PnCn+PnpCn}{ki1})^{c} + \frac{CLKBM1}{kt1}^{b} + 1} - dper \cdot per$$

$$\frac{d\,cry}{dt} = \frac{V2max\left(\frac{CLKBM1^{3} \cdot d}{kt2^{3}} + 1\right)}{\left(\frac{REVERBn}{ki21}^{f1} + 1\right)\left(\frac{CLKBM1^{3}}{kt2^{3}} + \frac{CLKBM1}{kt2}^{e}\left(\frac{PnCn+PnpCn}{ki2}\right)^{f} + 1\right)} - cry \cdot dcry$$

$$\frac{d\,Cc}{dt} = -Cc \cdot Pc \cdot kfPcCc - Cc \cdot Pcp \cdot kfPcpCc - Cc \cdot dCc + PcCc \cdot kdPcCc + PcpCc \cdot kdPcpCc + cry \cdot kp2$$

$$\frac{d\,Pc}{dt} = -Cc \cdot Pc \cdot kfPcCc - Pc \cdot dPc - Pc \cdot kphPc + PcCc \cdot kdPcCc + Pcp \cdot kdphPcp + kp1 \cdot per$$

$$\frac{d\,Pcp}{dt} = -Cc \cdot Pcp \cdot kfPcpCc + Pc \cdot kphPc - Pcp \cdot dPcp - Pcp \cdot kdphPcp + PcpCc \cdot kdPcpCc$$

$$\frac{d\,PcpCc}{dt} = Cc \cdot Pcp \cdot kfPcpCc - PcpCc \cdot dPcpCc - PcpCc \cdot kdPcpCc - PcpCc \cdot kiPcpCc + PnpCn \cdot kePnpCn$$

$$\frac{d\,PcCc}{dt} = Cc \cdot Pc \cdot kfPcCc - PcCc \cdot dPcCc - PcCc \cdot kdPcCc - PcCc \cdot kiPcCc + PnCn \cdot kePnCn$$

$$\frac{d\,PnpCn}{dt} = PcpCc \cdot kiPcpCc - PnpCn \cdot dPnpCn - PnpCn \cdot kePnpCn$$

$$\frac{d\,PnCn}{dt} = PcCc \cdot kiPcCc - PnCn \cdot dPnCn - PnCn \cdot kePnCn$$

Model 5: Model from Leloup & Goldbeter, 2003 [2].

$$\frac{d\,MP}{dt} = -MP \cdot kdmp - \frac{MP \cdot vmP}{KmP + MP} + \frac{vsP \cdot BN^n}{BN^n + KAP^n}$$
$$\frac{d\,MC}{dt} = -MC \cdot kdmc - \frac{MC \cdot vmC}{KmC + MC} + \frac{vsC \cdot BN^n}{BN^n + KAC^n}$$
$$\frac{d\,MB}{dt} = -MB \cdot kdmb - \frac{MB \cdot vmB}{KmB + MB} + \frac{vsB \cdot KIB^m}{BN^m + KIB^m}$$
$$\frac{d\,PC}{dt} = -CC \cdot PC \cdot k3 + MP \cdot ksP - \frac{PC \cdot V1P}{Kp + PC} - PC \cdot kdn + PCC \cdot k4 + \frac{PCP \cdot V2P}{Kdp + PCP}$$
$$\frac{d\,CC}{dt} = -CC \cdot PC \cdot k3 - \frac{CC \cdot V1C}{CC + Kp} - CC \cdot kdnc + \frac{CCP \cdot V2C}{CCP + Kdp} + MC \cdot ksC + PCC \cdot k4$$
$$\frac{d\,PCP}{dt} = \frac{PC \cdot V1P}{Kp + PC} - \frac{PCP \cdot V2P}{Kdp + PCP} - PCP \cdot kdn - \frac{PCP \cdot vdPC}{Kdp + PCP}$$
$$\frac{d\,CCP}{dt} = \frac{CC \cdot V1C}{CC + Kp} - \frac{CCP \cdot V2C}{CCP + Kdp} - CCP \cdot kdn - \frac{CCP \cdot vdCC}{CCP + Kd}$$
$$\frac{d\,PCC}{dt} = CC \cdot PC \cdot k3 - \frac{PCC \cdot V1PC}{Kp + PCC} - PCC \cdot k1 - PCC \cdot k4 - PCC \cdot kdn + \frac{PCCP \cdot V2PC}{Kdp + PCCP} + PCN \cdot k2$$
$$\frac{d\,PCN}{dt} = -BN \cdot PCN \cdot k7 + IN \cdot k8 + PCC \cdot k1 - \frac{PCN \cdot V3PC}{Kp + PCN} - PCN \cdot k2 - PCN \cdot kdn + \frac{PCNP \cdot V4PC}{Kdp + PCNP}$$
$$\frac{d\,PCCP}{dt} = \frac{PCC \cdot V1PC}{Kp + PCC} - \frac{PCCP \cdot V2PC}{Kdp + PCCP} - PCCP \cdot kdn - \frac{PCCP \cdot vdPCC}{Kd + PCCP}$$
$$\frac{d\,PCNP}{dt} = \frac{PCN \cdot V3PC}{Kp + PCN} - \frac{PCNP \cdot V4PC}{Kdp + PCNP} - PCNP \cdot kdn - \frac{PCNP \cdot vdPCN}{Kd + PCNP}$$
$$\frac{d\,BC}{dt} = \frac{-BC \cdot V1B}{BC + Kp} - BC \cdot k5 - BC \cdot kdn + \frac{BCP \cdot V2B}{BCP + Kdp} + BN \cdot k6 + MB \cdot ksB$$
$$\frac{d\,BCP}{dt} = \frac{BC \cdot V1B}{BC + Kp} - \frac{BCP \cdot V2B}{BCP + Kdp} - BCP \cdot kdn - \frac{BCP \cdot vdBC}{BCP + Kd}$$
$$\frac{d\,BN}{dt} = BC \cdot k5 - BN \cdot PCN \cdot k7 - \frac{BN \cdot V3B}{BN + Kp} - BN \cdot k6 - BN \cdot kdn + \frac{BNP \cdot V4B}{BNP + Kdp} + IN \cdot k8$$
$$\frac{d\,BNP}{dt} = \frac{BN \cdot V3B}{BN + Kp} - \frac{BNP \cdot V4B}{BNP + Kdp} - BNP \cdot kdn - \frac{BNP \cdot vdBN}{BNP + Kd}$$
$$\frac{d\,IN}{dt} = BN \cdot PCN \cdot k7 - IN \cdot k8 - IN \cdot kdn - \frac{IN \cdot vdIN}{IN + Kd}$$

4.2.4 Selection of parameters for FBXL3 and CKI

For FBXL3-CRY, parameters that determined the degradation rate of CRY (or CRY containing complexes) were considered to be the most likely candidates. Michealis-Menten degradation parameters were omitted from figure 4.7 since perturbations to such parameters are not easily attributable to changes in FBXL3 binding affinity. In the model presented in [2], CRY is degraded through a series of phosphorylation events, and these parameters were considered as representative of the rate of progression toward ubiquitination of CRY. The forward phosphorylation rates of CRY and nuclear PER-CRY complex were therefore also considered. For CKI-PER, we considered rates that determined the degradation rate and nuclear import rate of PER. Michealis-Menten parameters were not included, similar to FBXL3-CRY. With CRY being the main repressor of E box transcription [84], the degradation rates of PER-CRY complex were not considered as potential mechanisms of CKI. In the models of [2] and [1], the nuclear entry of PER-CRY requires two independent steps: the formation of the PER-CRY complex and the subsequent import of the complex. Therefore, the forward reaction rates of each of these steps were included.

4.2.5 Numerical experiments

Numerical parameter inhibitions were performed by recalculating the limit cycle trajectory for each new parameter set to a tolerance of 10^{-8}, using computational methods described previously [31].

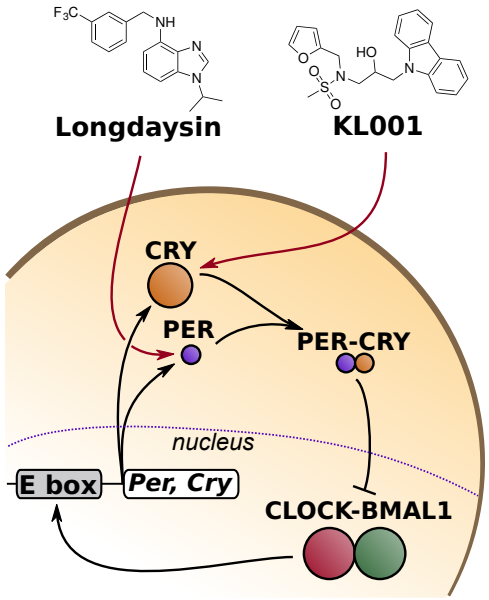


Figure 4.2: Small molecule targets. Schematic of the core circadian feedback loop, with the targets of small molecule modulators longdaysin (CKI inhibitor) and KL001 (inhibitor of FBXL3-dependent CRY degradation) shown. The size of each molecule is representative of relative concentration. Figure reprinted from [96].

4.3 Results

4.3.1 Opposite amplitude effects of longdaysin and KL001

To better understand both the CKI-PER and FBXL3-CRY pathways, we first studied the effects of small molecule compounds longdaysin and KL001, which cause stabilization of PER and CRY, respectively [49, 71] (figure 4.2). We used *Bmal1*- and *Per2-dLuc* as circadian reporters, which represent different loops of the core clock mechanism and show circadian luminescence rhythms with mutually opposite phase. Time-course data on circadian reporter expression under increasing concentrations of longdaysin and KL001 [49] was analyzed for period and amplitude change (figure 4.3). Longdaysin caused dose-dependent increases in period and detrended am-

plitude to ≈ 50% of control values in both *Bmal1-* and *Per2-dLuc* reporter cells. In contrast, KL001 induced a simultaneous increase in period and strong reduction in amplitude. Altering the activity of CKI-PER and FBXL3-CRY is therefore differentiated by an opposite amplitude response.

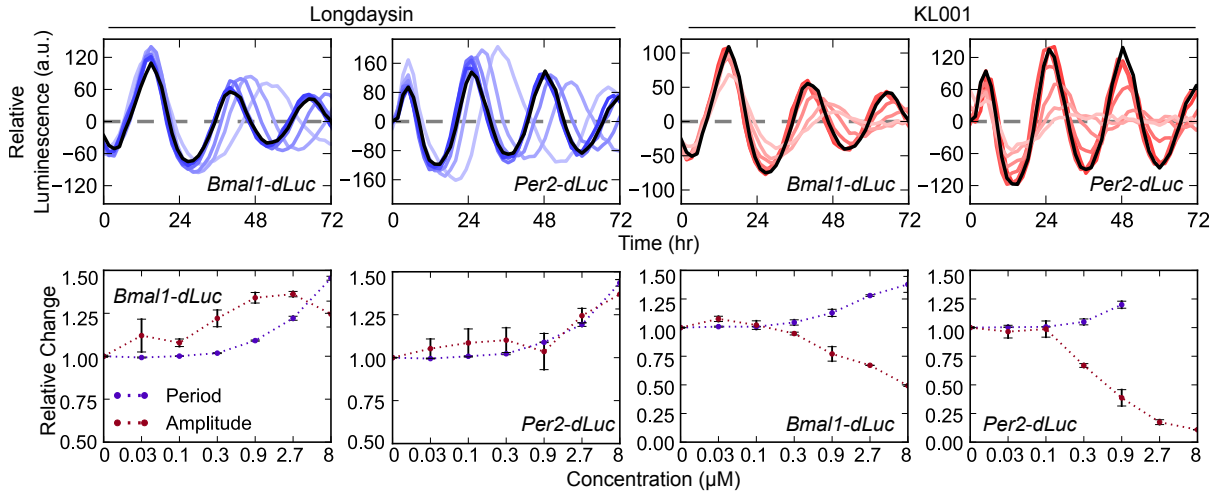


Figure 4.3: Different amplitude effect of small molecule circadian modulators targeting CKI-PER and FBXL3-CRY. (top) Detrended luminescence profiles (first 72 hours, mean of two independent replications) obtained from U2OS reporter cells with increasing concentration of longdaysin and KL001. Black profiles indicate control conditions (0 μM), lighter colors indicate higher concentrations of small molecule (from 0.03 to 8 μM). (bottom) Relative change in period and amplitude of the results in shown on top. Figure reprinted from [96].

In order to evaluate potential cross-interactions between CKI-PER and FXBL3-CRY pathways, the effect of longdaysin and KL001 on PER and CRY abundance was characterized by using PER1-LUC and CRY1-LUC reporters constitutively expressed in HEK293 cells (figure 4.4). Levels of PER1-LUC were increased only in the presence of longdaysin, while levels of CRY1-LUC were similarly increased only in the presence of KL001. These results suggest a lack of crossover between the two mechanisms.

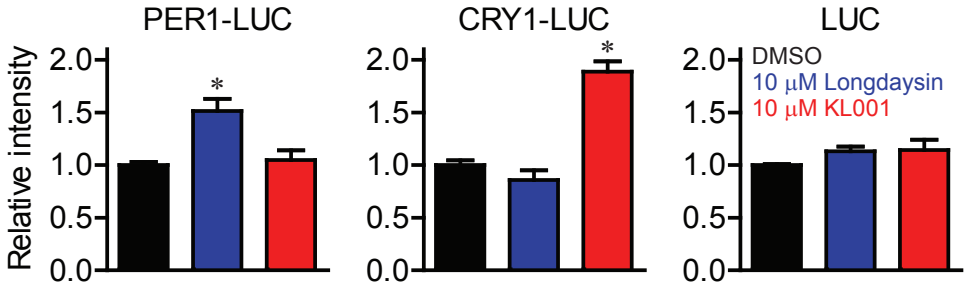


Figure 4.4: Effects of longdaysin and KL001 on PER1-LUC and CRY1-LUC abundance. HEK293 stable cell lines expressing PER1-LUC, CRY1-LUC, or LUC from a constitutive promoter were treated with DMSO, 10 μM longdaysin, or 10 μM KL001 for 24 h, and luminescence was measured. The luminescence intensity relative to DMSO control is shown as mean $\pm$ SEM ($n = 4 - 8$). * $P < 0.001$ compared to DMSO (one-way ANOVA with Dunnett's *post hoc* test).

4.3.2 Main period-determining perturbations

We next used *in silico* modeling to gain mechanistic insight into CKI-PER and FBXL3-CRY mediated circadian regulation. In chapter 2, we describe the connection between inhibition of FBXL3-dependent CRY degradation and period change [49]: increasing the stability of nuclear CRY results in longer transcriptional repression and increased period length. However, while CKI has been linked to modulating PER stability and nuclear entry, it remained unclear which perturbation dominates the period effect, and whether these processes are sufficient to separate the effects of CKI and FBXL3.

While mathematical models can be used to determine if a hypothetical mechanism is feasible, many approximations of the experimental system must be made. To generate predictions that are consistent across slight differences in model assumptions, we chose three mathematical models from the literature based on their moderate size and similar scope [1, 2, 49]. The models included, at a minimum, the expression and nuclear entry mechanisms of PER and CRY. We considered the formation of the PER-CRY heterodimer as a key step in nuclear entry, which is supported by the fact

that, to the best of our knowledge, all circadian models that consider both PER and CRY employ this kinetic assumption [1, 2, 43, 44, 49].

Dynamic models of genetic regulatory networks are typically comprised of a set of reaction equations and their associated kinetic parameters, which are chosen such that the model best fits the available experimental data. These kinetic parameters play a large role in determining the model's predictions. However, the sparsity of available data and high dimensionality of the parameter sets can result in a lack of confidence in the value of any particular fit [97]. In this report, we demonstrate that our predictions are independent of parameterization by employing a bootstrap identifiability analysis [81]. In a bootstrap method, multiple randomized *in silico* data sets are generated from experimental mRNA and protein concentration profiles [17]. For each of these trials, a new optimal parameter set is found by minimizing the difference between the model and the data. In this fashion, we are able to determine confidence intervals in model trajectories and output predictions as a function of the data quality. The state trajectories of the resulting 2000 parameter sets for each model are shown in figure 4.5, with reasonable agreement between models and experiment.

To identify parameters associated with PER and CRY protein activity that had the most reliable effect on period, we checked for parameter changes that shifted the period in a consistent direction across bootstrap trials. This search was achieved using a first-order period sensitivity analysis, which measures the derivative of the circadian period with respect to the value of the kinetic parameter (figure 4.6). To simplify analysis, we present only those rate parameters that are associated with experimentally supported mechanisms of CKI and FBXL3 in figure 4.7. We first tested parameters associated with potential FBXL3-CRY activity (figure 4.7A) to evaluate if our method matched the experimentally verified effect of KL001 [49]. Since CRY is the dominant repressor of CLOCK-BMAL1 [84], we attribute degradation rates of the

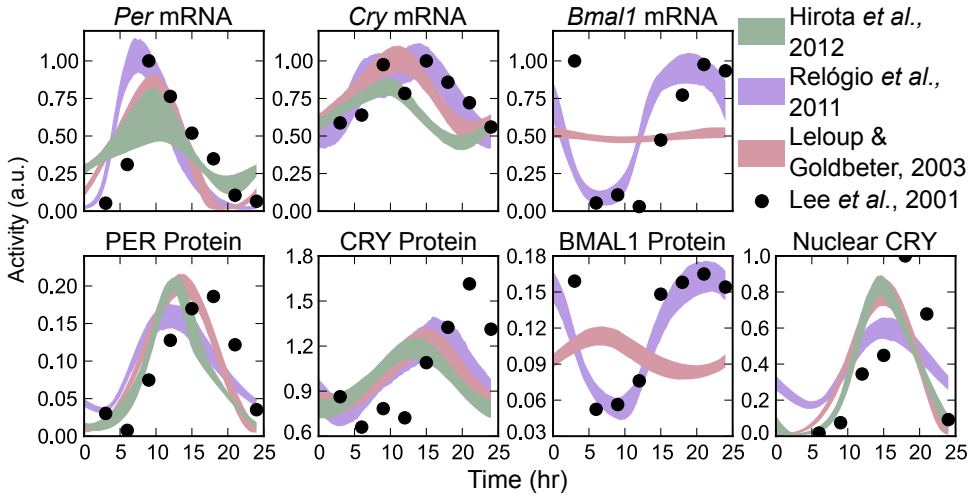


Figure 4.5: Fitted trajectories from bootstrap runs. Time series trajectories of the 2000 bootstrap trials for each model. Shaded regions indicate 95% confidence regions. The data were scaled to have a maximum value of 1, except for protein species, where relative values were important for clock stoichiometry. Figure reprinted from [96].

PER-CRY complex to be representative of CRY clearance rates. We found that only parameters governing nuclear CRY degradation show a period lengthening effect upon inhibition, while rates associated with cytoplasmic CRY degradation show period shortening effects. These results match with our previous assertion that period lengthening occurs via nuclear CRY stabilization [49]. Experimental evidence has also indicated cytoplasmic CRY2 stabilization may lead to period shortening [55], a result consistent with our mathematical predictions.

We next describe parameters potentially associated with CKI-dependent regulation of PER localization and stability (figure 4.7B). While the concentration (and therefore stability) of cytoplasmic PER plays a role in determining the nuclear entry rate, we treat the two rates as independent by comparing the period sensitivities of kinetic constants associated with each individual step. Since PER is rate-limiting in the formation of the PER-CRY complex [17], rates associated with complex formation or nuclear import were included in this analysis. Conversely, we did not include degra-

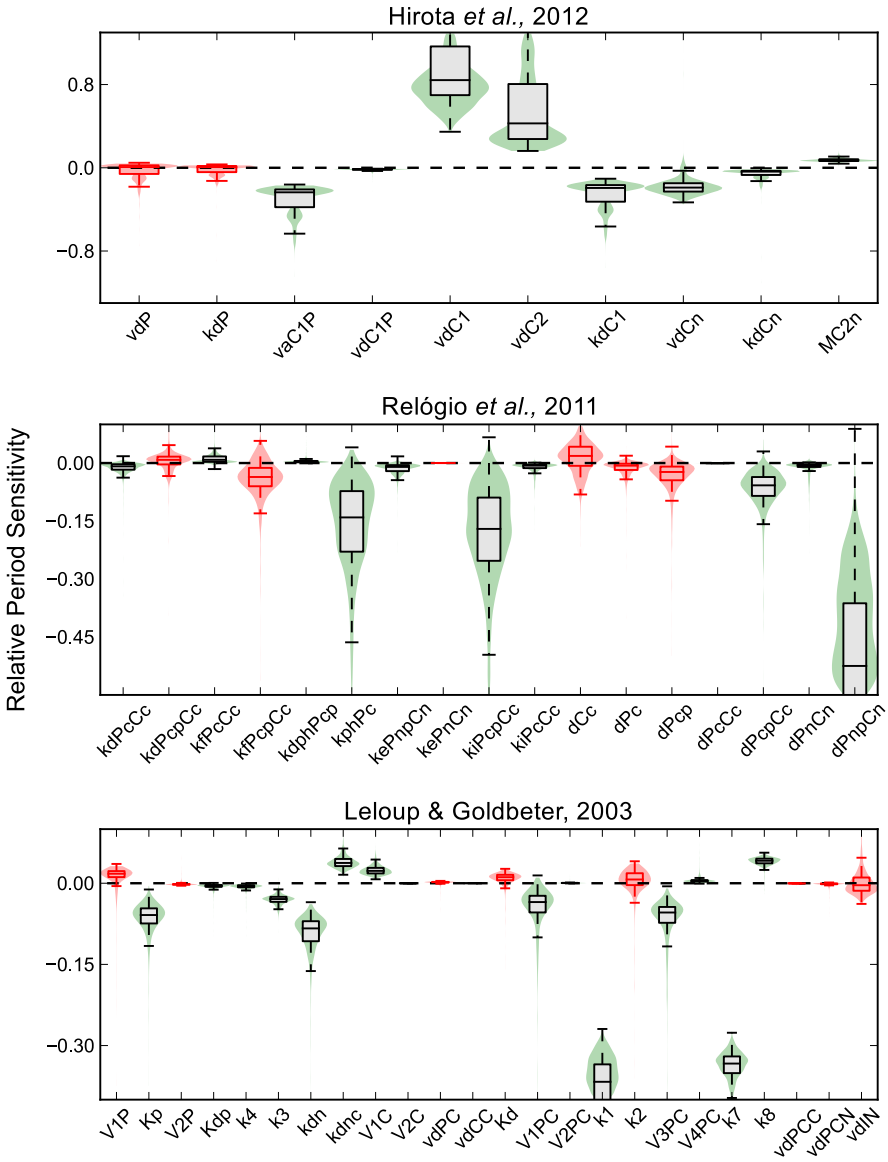


Figure 4.6: Bootstrap confidence intervals in relative period sensitivities. Violin plots for distributions in relative period sensitivities for parameters associated with PER and CRY proteins. Whiskers extend to the most extreme data point within $1.5x$ the inner quartile range. Distributions in which the 5^{th} and 95^{th} percentile lie on opposite sides of the x-axis are colored red and deemed non-identifiable. Figure reprinted from [96].

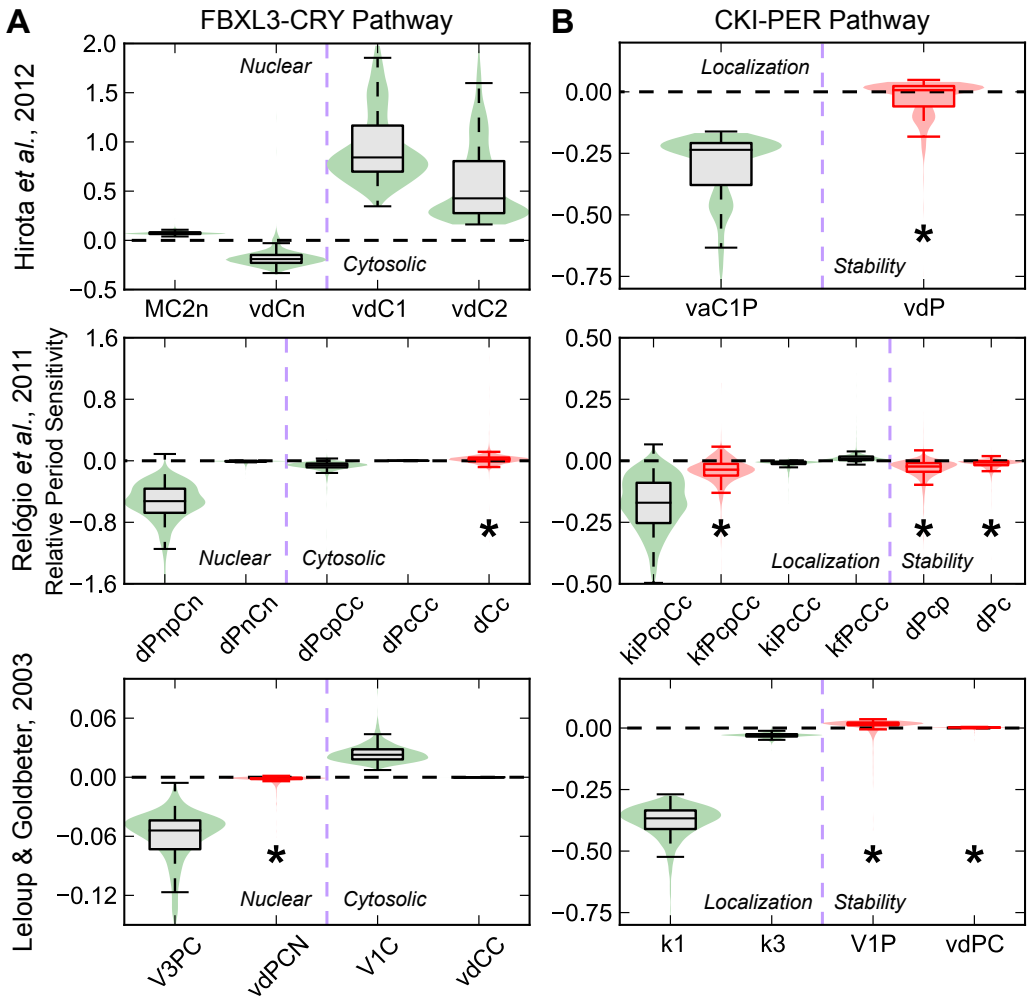


Figure 4.7: Bootstrap predictions of circadian actions of FBXL3-CRY and CKI-PER pathways. Violin plots of the relative period sensitivity of parameters associated with potential mechanisms for FBXL3-CRY (A) and CKI-PER (B) activity. A negative or positive period sensitivity indicate that the period of oscillation will increase or decrease when that rate is inhibited, respectively. Distributions that are not different than 0 with 95% confidence are colored red. Descriptions of the parameters shown are listed in tables 4.1-4.2. Figure reprinted from [96].

Table 4.1: Descriptions of the model parameters governing FBXL3-CRY

Model	Parameter	Description
[49]	MC2n	CRY2 nuclear multiplicative degradation coefficient
	vdCn	CRY1/2 nuclear degradation rate
	vdC1	CRY1 cytoplasmic degradation rate
	vdC2	CRY2 cytoplasmic degradation rate
[1]	dPnpCn	Phosphorylated PER-CRY complex nuclear degradation rate
	dPnCn	PER-CRY complex nuclear degradation rate
	dPcpCc	Phosphorylated PER-CRY complex cytoplasmic degradation rate
	dPcCc	PER-CRY complex cytoplasmic degradation rate
	dCc	CRY cytoplasmic degradation rate
[2]	V3PC	PER-CRY complex nuclear phosphorylation rate
	vdPCN	Phosphorylated PER-CRY complex nuclear degradation rate
	V1C	CRY cytoplasmic degradation rate
	vdCC	Phosphorylated CRY cytoplasmic degradation rate

Table 4.2: Descriptions of the model parameters governing CKI-PER

Model	Parameter	Description
[49]	vaC1P	PER-CRY complex nuclear entry rate
	vdP	PER cytoplasmic degradation rate
[1]	kiPcpCc	Phosphorylated PER-CRY complex nuclear entry rate
	kfPcpCc	Phosphorylated PER-CRY complex association rate
	kiPcCc	PER-CRY complex nuclear entry rate
	kfPcCc	PER-CRY complex association rate
	dPcp	Phosphorylated PER cytoplasmic degradation rate
	dPc	PER cytoplasmic degradation rate
[2]	k1	PER-CRY complex nuclear entry rate
	k3	PER-CRY complex association rate
	V1P	PER cytoplasmic phosphorylation rate
	vdPC	Phosphorylated PER cytoplasmic degradation rate

dation rates of PER-CRY nuclear repressive complex, since CRY alone is considered the main repressor. Surprisingly, even though both processes are related, model responses to altering PER stability were different from the responses to altering nuclear import directly. While it was previously hypothesized in models where PER acts as a direct repressor that the regulation of PER stability would determine the period [92, 95], our new assumptions revealed that parameters governing PER degradation showed only non-identifiable responses. Alternatively, inhibition of rates associated with the nuclear entry of the PER-CRY complex showed strong period lengthening effects. These results indicate that under our current understanding of clock kinetics, the regulation of nuclear import likely plays the prominent role in CKI-dependent period regulation.

4.3.3 Independent mechanisms of PER and CRY regulation

Using model 1 and the perturbations identified in figure 4.7, we first confirmed that inhibition of nuclear CRY degradation (vdCn) and PER-CRY nuclear import (vaC1P) reproduced the experimental period and amplitude effects of the small molecules KL001 and longdaysin, respectively (figure 4.8, compare with figure 4.3).

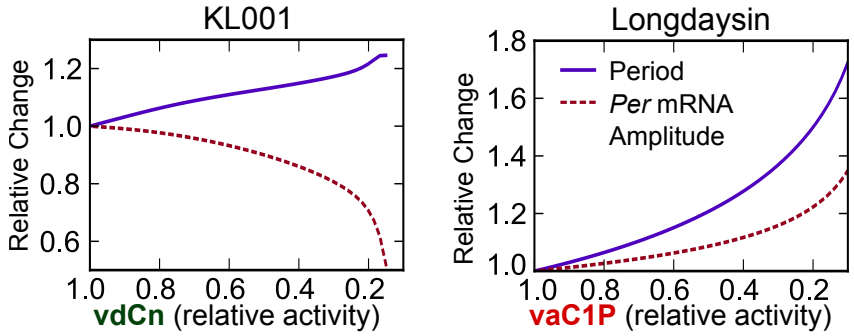


Figure 4.8: Predictions of KL001 and longdaysin results. *In silico* reproductions of the circadian reporter experiments in figure 4.3, using the predictions identified in figure 4.7. Figure reprinted from [96].

Comparison of the oscillatory profiles of *Per* mRNA and nuclear CRY protein (figure 4.9) revealed that inhibition of FBXL3-dependent CRY degradation caused lingering nuclear CRY to not be completely purged each cycle. This excess repressor during the accumulating phase of *Per* and *Cry* transcripts resulted in lower E box amplitudes, providing a likely explanation for the effect of KL001.

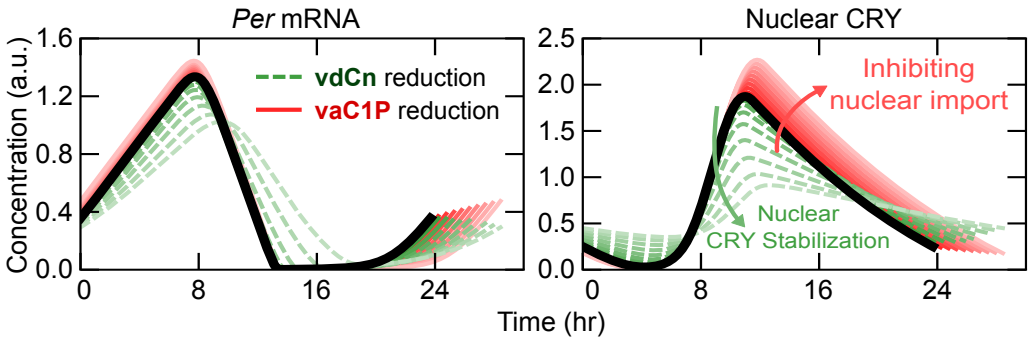


Figure 4.9: Comparison of the effects of KL001 and longdaysin. Parameter changes were normalized such that the period change was equal for each pair of perturbations (vdCn: 100% $\rightarrow$ 23%, vaC1P: 100% $\rightarrow$ 51%). Figure reprinted from [96].

In contrast, stabilization of cytoplasmic PER (lowering vdP) resulted in reduced transcriptional amplitude with minimal period effect (figure 4.10), consistent with experimental findings from the knockdown of β-TrCP, an F box protein responsible for PER degradation [98]. However, other experimental results have shown that down-regulation of β-TrCP leads to longer periods [86], suggesting that further modeling and experimental inquiry is needed on the role of β-TrCP in clock regulation. This period lengthening might be explained through β-TrCP-mediated stabilization of nuclear PER-CRY or by using alternative kinetic assumptions for the rate of PER-CRY binding.

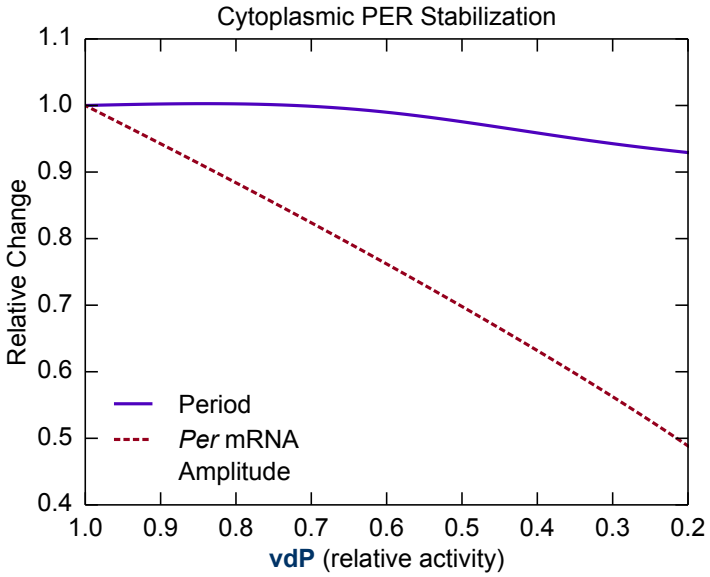


Figure 4.10: Effect of cytoplasmic PER stabilization on period and E box transcription amplitude. Relative period and peak-to-trough amplitude change in *Per* mRNA resulting from a reduction in the vdP parameter in the model from Hirota *et al.*, 2012. Figure reprinted from [96].

CKI likely regulates nuclear entry rather than PER stability

We further compared the effect of inhibiting PER degradation with inhibiting nuclear import on the oscillatory profile of key clock proteins (figure 4.11) to identify mechanistic differences between the two potential effects of CKI inhibition. Both perturbations increased cytoplasmic PER, suggesting the two mechanisms are difficult to distinguish experimentally. Direct stabilization of PER in the cytoplasm (lowering vdP) shortened the time delay between transcription and inactivation by accelerating the accumulation of cytoplasmic PER and nuclear PER-CRY. However, it also lengthened the repressive phase by increasing the total amount of PER-CRY that enters the nucleus. The net result of these opposing perturbations was little period change, as indicated by a faster accumulating phase and slower declining phase in nuclear CRY after vdP reduction (figure 4.11). In contrast, inhibiting PER-CRY nuclear entry (lowering vaC1P) caused additional free protein to build in the cytoplasm, delaying nu-

clear accumulation and ultimately increasing the total amount of nuclear PER-CRY. Since both of these trends work to increase period length, inhibiting PER-CRY nuclear entry resulted in longer cycles. Additionally, the longer cytoplasmic time delay resulted in increased transcription, yielding slightly higher amplitudes (figure 4.9) that closely match the experimental results of the small molecule longdaysin.

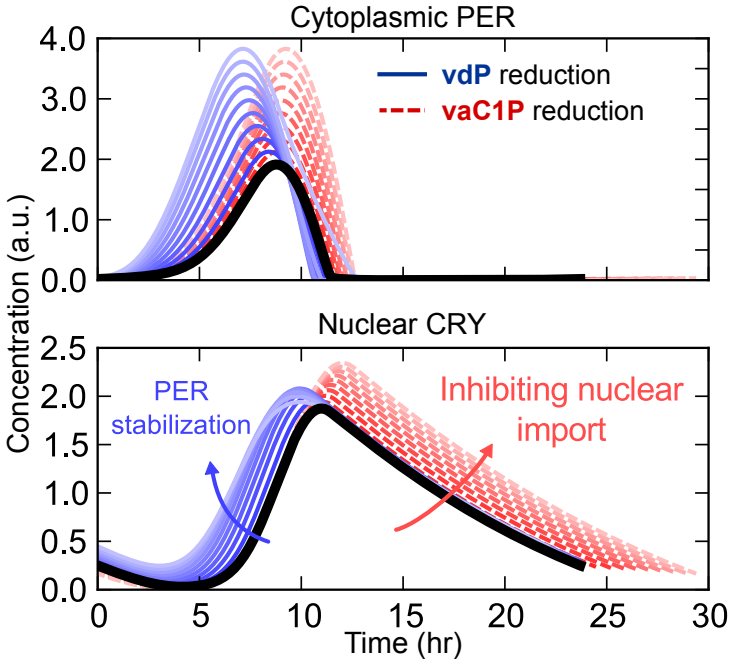


Figure 4.11: Comparison of two candidate mechanisms for CKI inhibition. Effects of increasing PER stabilization and nuclear import inhibition on the time profiles of cytoplasmic PER (top) and nuclear CRY (bottom). Parameter values were selected such that the amplitudes of cytoplasmic PER are equal at each level. Lighter colors indicate stronger perturbations (vdP: 100% $\rightarrow$ 22%, vaC1P: 100% $\rightarrow$ 45%), $t = 0$ is set to the onset of PER accumulation. Figure reprinted from [96].

Since CKI likely regulates both stability and subcellular localization of PER *in vivo*, we considered the effects of simultaneously lowering both PER cytoplasmic degradation and nuclear entry rates, as shown in figure 4.12. The loss of oscillations under extreme reduction of both parameters highlights an interesting role of CKI in conferring robustness to the circadian clock: since oscillations are lost when import of

the PER-CRY complex to the nucleus ceases to be rhythmic, CKI ensures lingering PER is purged from the cytoplasm by one pathway or another before E box transcription resumes. This importance has been proven experimentally, as disruption of CKI-mediated regulation leads to compromised circadian oscillations [99].

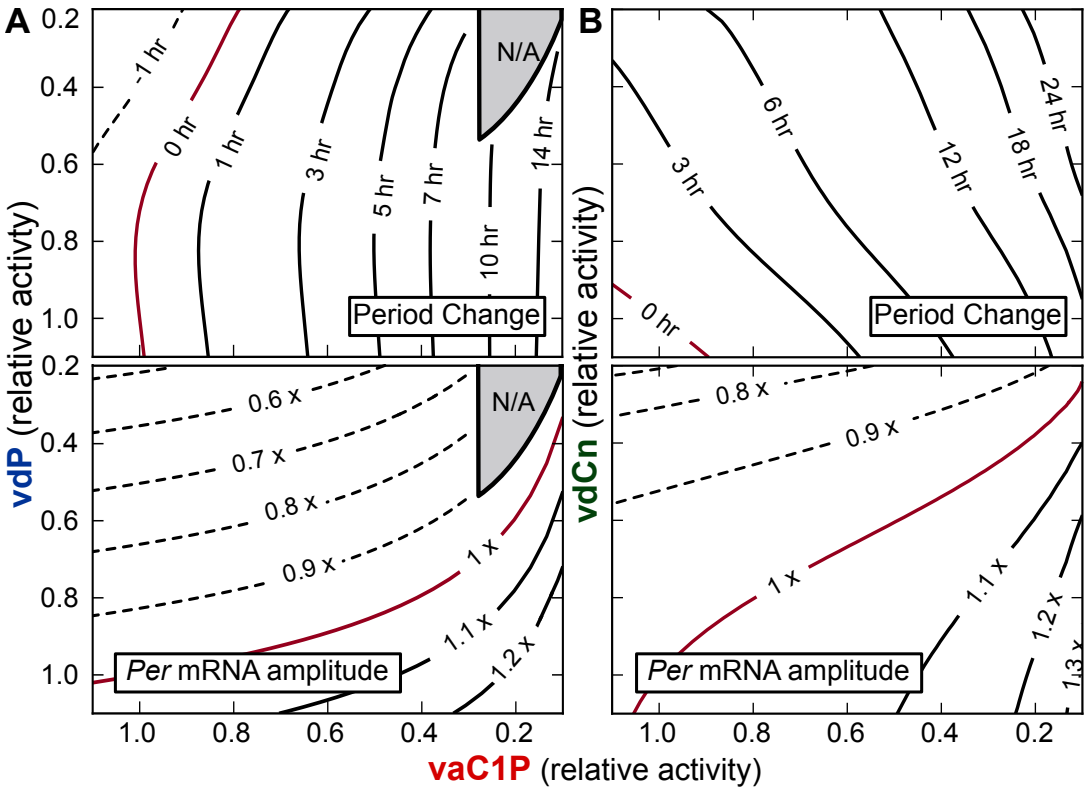


Figure 4.12: Independence of CKI-PER and FBXL3-CRY pathways. (**A**) Contour plots of period change (top) and **Per** mRNA amplitude (bottom) for simultaneous inhibition of both PER degradation (vdP) and nuclear import (vaC1P). The gray shaded region indicates loss of oscillations. (**B**) Period and amplitude change contour plots for varying both vdCn (CRY nuclear degradation rate) and vaC1P (PER-CRY nuclear import). Figure reprinted from [96].

Together, inhibition of CKI by longdaysin may increase the time required before PER-CRY can enter the nucleus to repress transcription, leading to a higher amplitude and longer period. In contrast, KL001 lengthens the period by stabilizing nuclear CRY, resulting in a longer time delay before transcription resumes and lower

amplitude from increased E box repression. PER regulation through CKI is therefore partitioned to the accumulating phase, controlling the speed and amount of PER-CRY complex that enters the nucleus. CRY regulation through FBXL3 is partitioned independently to the repressive phase, controlling the length of time until CLOCK-BMAL1-dependent transcription resumes (figure 4.13). This independence was reproduced *in silico* by the simultaneous reduction of nuclear CRY degradation and PER-CRY nuclear import (figure 4.12B).

Quantification of linear additivity

In order to quantify this biological independence, changes in period obtained from the simultaneous reduction in both FBXL3 and CKI activity were compared to the expected period change if both processes were linearly additive. R^2 values for figure 4.12B were calculated by comparing the actual results $f(x,y)$ to linearly additive surfaces $\hat{f}(x,y)$. The linearly additive surface takes the form $\hat{f}(x,y) = g(x) + h(y)$, where $g(x) = f(x,1)$ and $h(y) = f(1,y)$ (figure 4.14). The $f(\cdot)$ and $g(\cdot)$ functions were found by interpolating the x- and y- axis of figure 4.12B, respectively, in which a knockdown is only present on one of the two activities. The differences between the actual period changes and the period changes predicted by an additive model were quantified by using an R^2 value,

$$R^2 = 1 - \frac{\int_{x,y} f(x,y) - \hat{f}(x,y)\ dxdy}{\int_{x,y} f(x,y) - \mu\ dxdy},$$

in which

$$\mu = \frac{1}{\Delta x \Delta y} \int_{x,y} \hat{f}(x,y)\ dxdy.$$

R^2 calculations were restricted to the range $x, y \in [0.4, 1.1]$ to omit bifurcation regions. $R^2 = 0.924$ and 0.999 for period change and *Per* mRNA amplitude, respectively, in-

dicating net changes to clock kinetics from perturbations to both rates can be closely explained by simply adding together the individual effects from each pathway (figure 4.14).

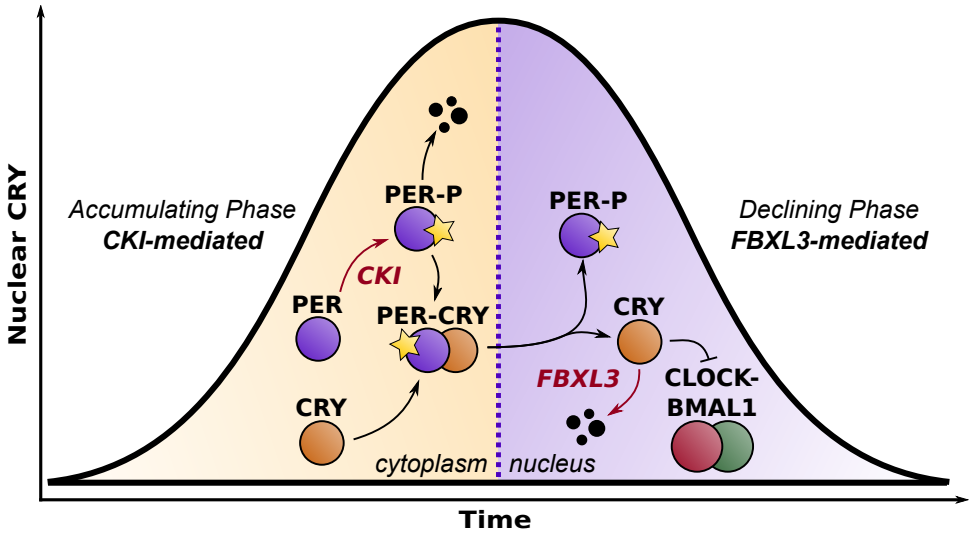


Figure 4.13: Spatiotemporal separation underlies independence of CKI-PER and FBXL3-CRY pathways. Longdaysin, acting through CKI, lengthens the accumulating phase of the circadian cycle, while KL001, acting through FBXL3, lengthens the declining phase. Figure reprinted from [96].

4.4 Discussion

An understanding of the interactions between posttranslational regulators is crucial for the further development of circadian pharmacological reagents, as efficient modulation of clock function will assuredly come from simultaneous perturbations to many connected species. In this chapter, we used circadian reporter cells together with mathematical modeling to provide mechanistic insight into the differences of CKI- and FBXL3- mediated posttranslational regulation of PER and CRY. As a result, we clarified a process by which CKI exerts control over the circadian period, demonstrated through both the hyperphosphorylating CKIϵ^{tau} mutant and small

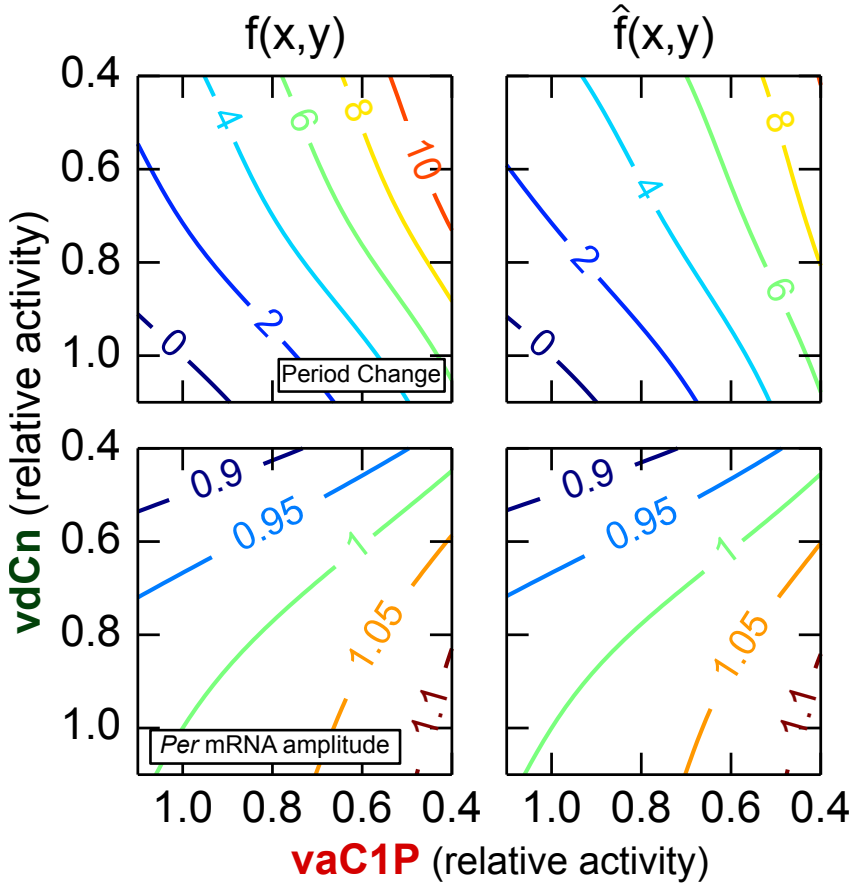


Figure 4.14: Calculation of R^2 values for figure 4.12B. Contour plots show changes to period and mRNA amplitude from inhibition of nuclear CRY degradation and PER nuclear import. Calculated values after applying both perturbations, $f(x,y)$, are compared to the linear sum of both individual perturbations, $\hat{f}(x,y)$. The similar profiles in both pairs of plots indicate cooperative and competitive effects are largely absent. $R^2 = 0.924$ and 0.999 for period change and *Per* mRNA amplitude, respectively. Figure reprinted from [96].

molecule CKI inhibitors, such as longdaysin. Future work will further explore the post-translational landscape surrounding PER and CRY, including the recently discovered FBXL21, the paralog of FBXL3, that is responsible for regulating CRY stability in the cytoplasm [100, 101]. In developing our predictions, we have used multiple models and parameterizations to ensure our mechanisms are consistent across many *in silico* realizations. These results reinforce the notion that computational modeling is essential in interpreting results in systems with complicated oscillatory feedback. Additionally, *in silico* analyses reveal hidden design principles of biological networks, as this work highlights the importance of the CKI family of kinases in conferring robustness to the circadian cycle.

4.4.1 Phase-dependent activity of small molecule modulators

In this chapter, it was shown that the posttranslational CKI and FBXL3 control separate time regimes of the core clock oscillation. While in the cultured cell experiments the small molecule modulators were continuously present, in pharmacological applications there would likely be dosed at specific times of day. Since their targets play key roles during specific circadian phases, it is likely that there would be a significant difference in small molecule effect depending on the time at which they are given. In the next chapter, we systematically analyze the dynamic consequences associated with *transient* applications of small molecule agonises, rather than dose-dependent continuous applications, as the former will likely have more significant clinical relevance.

Chapter 5

Amplitude metrics for cellular bioluminescence reporters[1]

5.1 Background

In the previous chapter, we describe used mathematical modeling to explain the systems-level effects of two pharmacological agents. In doing so, it was shown that each compound controlled reactions which primarily occur during specific times of day. Since clinical applications of circadian therapies will likely have a time at which they are administered, it is important to understand the effects of temporary perturbations on circadian amplitude. The effects of temporary perturbations are more complicated, however, since they may effect both the single-cell oscillator as well as the synchrony of cell populations as a whole. In this chapter, we develop the mathematical background required to understand and efficiently predict the response of a population of oscillators to a temporary change.

[1]Portions of this chapter are published in P. C. St. John, S. R. Taylor, J. H. Abel, and F. J. Doyle III, "Amplitude Metrics for Cellular Circadian Bioluminescence Reporters," *Biophys. J.*, vol. 107, pp. 2712– 2722, Dec. 2014.

5.1.1 Importance of circadian amplitude

While the suprachiasmatic nucleus (SCN) in the brain serves as the body's master pacemaker, cells found in peripheral tissues also oscillate in a circadian manner [102]. These peripheral clocks process systemic and SCN-mediated entraining cues, buffering against rhythmic changes in energy availability to maintain metabolic homeostasis [103]. The amplitude of circadian transcription is a relevant factor, and has been shown to play a critical role in phase resetting and entrainment [104, 105]. Recent studies have further highlighted the importance of high peripheral clock amplitudes in maintaining metabolic health. Mice lacking an intact clock have been shown to develop metabolic disease [12], while low amplitude clock oscillations, whether caused by diet [15] or age [106] have also been tied to metabolic disorders. An understanding of how clock amplitudes are regulated is therefore a topic of ongoing research with potential therapeutic applications [96]. Unlike the network of oscillators in the SCN, in which intercellular coupling maintains robust amplitudes even in the absence of external cues, peripheral oscillators are thought to lack a direct mechanism to spontaneously synchronize [25]. As a result, populations of peripheral clocks are likely synchronized by common external cues, with stochastic effects and cell heterogeneity driving entrained populations gradually toward desynchrony.

5.1.2 Amplitude at the single-cell and population level

The development of immortalized peripheral oscillator cell lines with bioluminescent reporters has allowed high-throughput analysis of the responses of circadian rhythms to genetic and pharmacological manipulation [71, 107]. In addition to providing experimental tractability, these systems provide detailed information on the amplitude of oscillations in gene transcription, a measure often lacking from earlier experiments

using wheel-running activity. As a result, these cell lines have proven useful in studying core clock connectivity and stoichiometry [108]. However, since *in vitro* experiments typically measure entire cultures of cells, data collected at the population-level can obscure the response of the clock at the scale of the gene-regulatory network. Even when individual cells can be recorded, stochastic noise hinders accurate amplitude determination. As a result, two studies using similar perturbations to understand the mechanism of light-induced amplitude reduction reached differing conclusions, in which either single-cell amplitude reduction or population-level desynchrony was identified as the dominant factor [109, 110].

5.1.3 Response functions

Response functions are general techniques that maps a change in an output variable to a temporary change in parameters [111]. For instance, the phase response curve (PRC) has been used to characterize the entrainment behavior of both experimental and mathematical systems [112, 113, 114], and in analyzing the synchrony of populations of oscillators [115]. Accurate and efficient numerical routines for finding infinitesimal PRCs have therefore been developed [30, 32, 97]. In addition to changes in phase, phase-dependent changes in amplitude are an important factor in understanding circadian rhythms. Amplitude response curves (ARCs) were first used in the clock literature to understand the effects of light pulses on simple phase-amplitude models, and were useful in predicting phase singularity behavior [116]. ARCs have also been used to characterize perturbations to groups of oscillators through desynchrony [110, 117], at the level of the single oscillator [118, 119], and even in studying the effect of entrainment phase on SCN rhythm amplitude [120]. However, previous definitions of the ARC are inconsistent between studies and do not simultaneously consider amplitude effects at the single-cell and population level.

The choice of modeling framework dictates the type of amplitude response which can be predicted. Ordinary differential equation (ODE) models of gene regulation are capable of describing the amplitude and phase-resetting behavior of single cells, but fail to capture the collective dynamics of a population of oscillators. Likewise, phase-only models correctly capture the change in synchrony of a population, but do not capture fluctuations in amplitude of individual oscillators. Explicit stochastic simulations of populations of cells are capable of realistically capturing both single-cell and population-level effects, and have been successfully used to understand the response of coupled oscillators to external VIP perturbation [121]. However, these methods are computationally expensive, and cannot be used for an analytical understanding of amplitude response.

In this chapter, we describe approaches to quantify amplitude change in a population of non-interacting oscillators. By exploiting the independence of each oscillator, we derive computationally efficient methods to approximate the mean dynamics of full stochastic simulations. Additionally, our method allows the calculation of ARCs at both the single-cell and population level, allowing the behavior of the system to be quickly profiled. Specifically, we use ODE models to describe the transient amplitude response at the single-cell level, coupled with a phase probability density function to describe population-level dynamics. Following a perturbation with a finite duration, a limit cycle oscillator undergoes a transient change in amplitude. When the mean expression level of a population of oscillators is also considered, an additional change in amplitude is incurred due to the change in synchrony of the population, which persists until synchrony is changed by subsequent perturbations. We therefore observe a separation in timescales between the effects on clock output mediated at the single-cell level and those mediated by population synchrony, allowing the source of an amplitude change to be qualitatively inferred by inspection of bioluminescence

data. Characterizing the mechanisms and consequences of both types of amplitude regulation will be important in understanding how peripheral amplitudes are maintained, and may lead to the design of pharmacological or behavioral strategies to boost circadian amplitudes.

5.2 Methods

5.2.1 Perturbations to limit cycle systems

In this chapter we restrict our analysis to temporary perturbations capable of entraining an oscillatory system, excluding permanent parameter changes arising, for instance, from genetic knockout experiments. The most simple entraining perturbation involves adding or removing components to a limit cycle system, resulting in a perturbed trajectory $x(t)$. Here the initial conditions are determined by the strength of the perturbation and the phase at which it is applied:

$$x(0) := x^{\gamma}(\theta_0) + \Delta x(0). \tag{5.1}$$

This trajectory evolves according to equation (1.1), eventually returning to x^{γ}. It is useful to express this trajectory by the deviation from the limit cycle:

$$\Delta x(\tilde{t}) = x(\tilde{t}) - x^{\gamma}(\tilde{t} + \theta_0). \tag{5.2}$$

In addition to perturbations directly to the state of the system, oscillators can also be perturbed by temporary changes to the parameters. For a parameter pulse Δp of duration $\tilde{d}$ (in radians) that ends at θ_0, the oscillator deviates from the limit cycle

trajectory according to:

$$x_{\tilde{d}}(-\tilde{d}) = x^{\gamma}(\theta_0 - \tilde{d}) \tag{5.3}$$

$$\frac{dx_{\tilde{d}}}{d\tilde{t}} = \tilde{f}(x_{\tilde{d}}(\tilde{t}), p + \Delta p). \tag{5.4}$$

The pulse trajectory $x_{\tilde{d}}(\tilde{t})$ is then integrated from $\tilde{t} = -\tilde{d} \rightarrow 0$, at which point the pulse is removed. A temporary parameter pulse is therefore equivalent to a state perturbation, but with the perturbation at $t = 0$ defined by:

$$\Delta x(0) = x_{\tilde{d}}(0) - x^{\gamma}(\theta_0). \tag{5.5}$$

A schematic depicting an example perturbed and reference trajectory is shown in figure 5.1.

5.2.2 Phase-diffusion model

Large populations of oscillators are typically described using phase-only models [123], in which the state of each oscillator is represented only by its phase, θ. The synchrony of the population can be modeled using a probability density function $p(\theta, \tilde{t})$ that describes the probability of finding an oscillator at each phase [124]. The usefulness of probability density functions in describing the phase and amplitude responses of populations of circadian cells has been previously shown [110]. As with all probability density functions,

$$\int_0^{2\pi} p(\theta, \tilde{t})\, d\theta = 1. \tag{5.6}$$

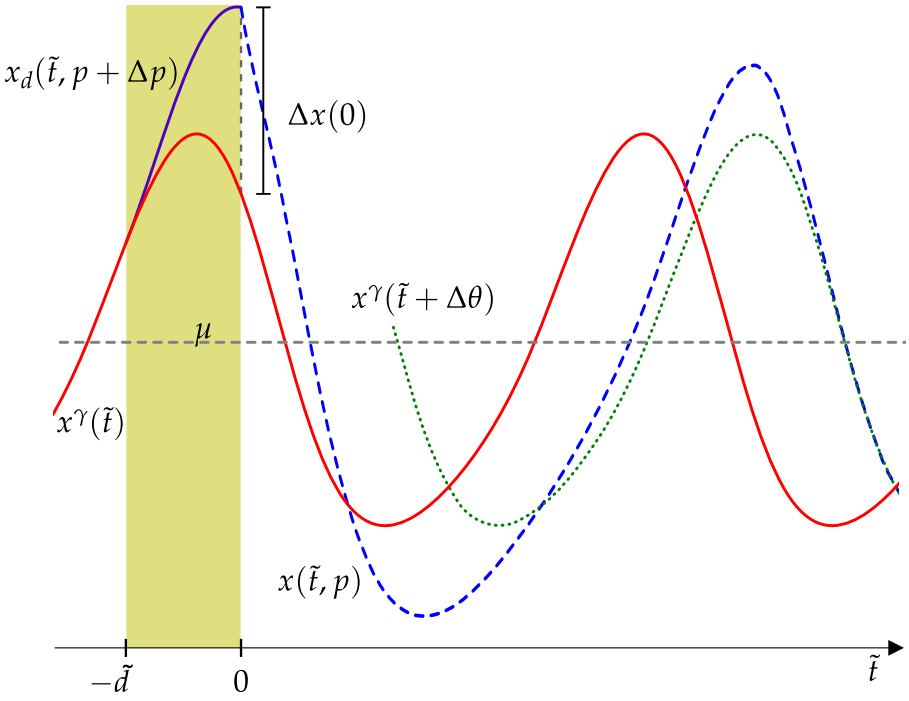


Figure 5.1: Schematic showing trajectories used in the calculation of single-cell phase and amplitude change. A perturbation $x(\tilde{t}, p)$, blue, from the limit cycle solution $x^\gamma(\tilde{t}, p)$, red, ultimately returns to the limit cycle with a phase shift $x^\gamma(\tilde{t} + \Delta\theta, p)$, green. Figure reprinted from [122].

The shape of $p(\theta, \tilde{t})$ changes as the cells advance in time. Stochastic effects cause the population to gradually desynchronize as slight cycle-to-cycle deviations are propagated throughout the population [125]. For a infinite population of oscillators, these effects are well-described by a Fokker-Plank equation [126]:

$$\frac{\partial p}{\partial \tilde{t}} = \frac{\partial p}{\partial \theta} + d\frac{\partial^2 p}{\partial \theta^2}. \tag{5.7}$$

Due to the rescaling of t, the mean period of the population is 2π. Here, the $\partial p/\partial \theta$ term, analogous to convection, describes the mean oscillatory period, while the $\partial^2 p/\partial \theta^2$ term describes the diffusion of phases across $[0, 2\pi)$. The phase diffusivity parameter d (in units of inverse radians) describes the speed with which the population desynchronizes and can be fit to experimental data [127]. equation (5.7) has periodic boundary conditions, with initial condition $\phi(\theta)$ as the phase population at $t = 0$:

$$\text{BCs:} \quad p(0, \tilde{t}) = p(2\pi, \tilde{t}) \tag{5.8}$$

$$\frac{\partial p}{\partial \theta}(0, \tilde{t}) = \frac{\partial p}{\partial \theta}(2\pi, \tilde{t}) \tag{5.9}$$

$$\text{IC:} \quad p(\theta, 0) = \phi(\theta). \tag{5.10}$$

The solution of equations (5.7-5.10) is well-characterized, with $p(\theta, \tilde{t})$ evolving in time as the convolution of the initial conditions with a wrapped normal distribution with mean $\tilde{t}$ and standard deviation $\sqrt{2d\tilde{t}}$ [128]:

$$p(\theta, \tilde{t}) = \phi(\theta) * \mathcal{WN}(\theta; \tilde{t}, \sqrt{2\,d\,\tilde{t}}), \tag{5.11}$$

in which the wrapped normal distribution [129] is defined as

$$\mathcal{WN}(\theta;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \exp\left[\frac{-(\theta-\mu+2\pi k)^2}{2\sigma^2}\right]. \tag{5.12}$$

Since the convolution of two normal distributions is also a normal distribution, it is efficient when possible to describe $\phi(\theta)$ as a normal distribution with mean μ_0 and standard deviation σ_0, such that $p(\theta,\tilde{t})$ can be found analytically through

$$p(\theta,\tilde{t}) = \mathcal{WN}(\theta;\mu_0+\tilde{t},\sqrt{\sigma_0^2+2d^2\tilde{t}^2}). \tag{5.13}$$

5.2.3 Numerical simulations

ODE models were simulated in Python, using the computer algebra package CasADi [77]. Stochastic simulations were performed using the StochKit2 package [130]. The codes used to generate the figures can be found online at `https://github.com/pstjohn/circadian-amplitude-metrics`

5.3 Results

A signal to a population of limit cycle oscillators can affect amplitude in two ways. First, individual cells are perturbed from their limit cycles, and exhibit transient dynamics before settling back to steady state amplitudes. Secondly, the phases of the population are changed, resulting in a permanent change in population synchrony. Differences between these two types of amplitude change are summarized in figure 5.2. Therefore, in order to derive continuous approximations to the dynamics of a large population of non-interacting oscillators, we first demonstrate how ARCs can be found at both the single-cell and population level.

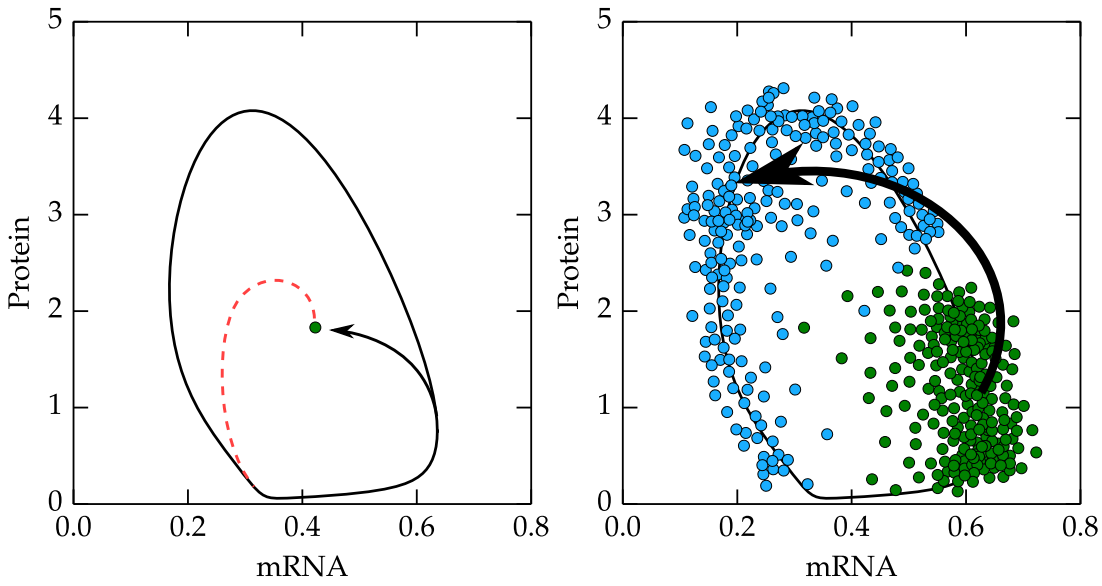


Figure 5.2: Differences between population and single-cell mediated amplitude change. Amplitude change at the single-cell level (left) is temporary, mediated by the dynamics of the limit cycle system. At the population-level (right), many limit cycle oscillators are averaged together to find the total expression level. The amplitude is therefore changed by the degree of synchrony of the population, which is often changed by perturbations. Figure reprinted from [122].

5.3.1 Definition of an amplitude metric at the single-cell level

Following a perturbation to a single cell, the path that the perturbed trajectory $x(t)$ takes in returning to the limit cycle will have a different amplitude than the unperturbed trajectory. While such an amplitude change will only have a finite duration, it plays an important role when perturbations are repeatedly received by the clock, such as a peripheral oscillator entrained to daily metabolic stimuli. To define an amplitude change metric for such a case, we compare a perturbed trajectory $x(t)$ to a phase-shifted limit cycle reference $y(t)$, for which $x(t) \to y(t)$ for sufficiently long times. Since $x(t)$ approaches the reference as $t \to \infty$, the means of both trajectories are equal and can be calculated by

$$\mu \coloneqq \int_0^{2\pi} \frac{x^\gamma(\theta)}{2\pi}\, d\theta. \tag{5.14}$$

The amplitude change metric is defined as

$$\begin{aligned} \Delta A(x(t), y(t)) &\coloneqq \int_0^\infty (x(t)-\mu)^2 - (y(t)-\mu)^2\, dt \\ &= \int_0^\infty h(t)\, dt. \end{aligned} \tag{5.15}$$

This amplitude metric was chosen over alternatives, such as peak-trough distance, for its analytical tractability. The integrand in equation (5.15), abbreviated by $h(t)$, compares the variances of the reference and perturbed trajectories. When $h(t) > 0$, the trajectory is further from the mean than the reference, and similarly when $h(t) < 0$ the trajectory is closer to the mean than the reference. Thus the overall amplitude change can be calculated by integrating $h(t)$ until the two trajectories converge, returning an amplitude value for each state variable. Amplitude change for a two-dimensional oscillator is easy to visualize graphically. In figure 5.3, the same state

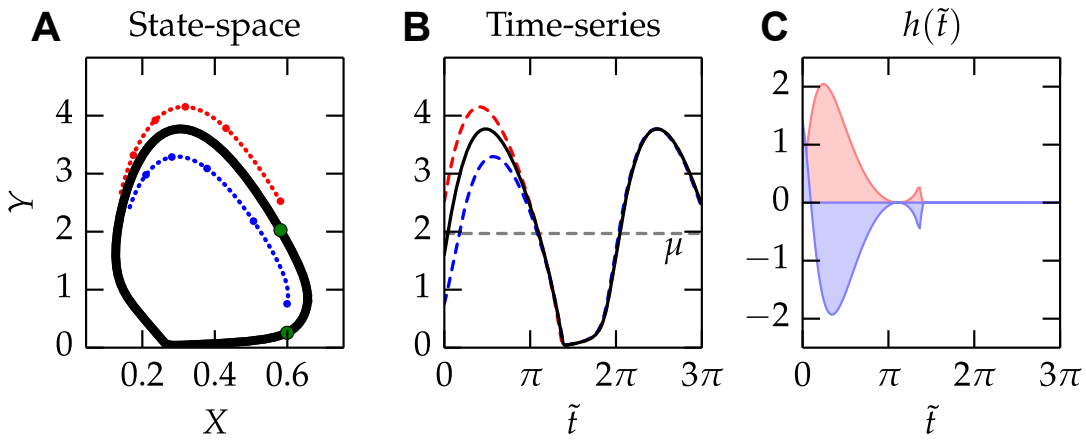


Figure 5.3: Amplitude metrics at the single-cell level measure transient deviations from the limit cycle. The same perturbation applied at two different phases can result in opposite amplitude effects. The state-space representation (**B**) reveals the path both perturbations take to return to the limit cycle. The time-series representation (**C**) shows how the perturbation in blue results in an amplitude decrease, while the one in red results in an amplitude increase. (**D**) This amplitude change is quantified by integrating $h(\tilde{t})$, the difference in variance from each solution to the limit cycle as defined in equation (5.15), from $t = 0 \rightarrow \infty$. Here the shaded region indicates the area under the curve, ΔA, for the perturbations shown in red and blue. Model adapted from [3]. Figure reprinted from [122].

perturbation $\Delta x(0)$ applied at two different phases results in opposite amplitude changes, depending on whether the perturbation shifts the trajectory to the interior of the limit cycle (reduced amplitudes) or to the outside the limit cycle (increased amplitudes). Trajectories for the second state variable, $Y(\tilde{t})$, and the corresponding integrand for the amplitude change equation, $h(\tilde{t})$, demonstrate how this transient change is quantified.

5.3.2 Single-cell amplitude response curves

Similar to the PRC, we denote the phase-dependent amplitude change following a perturbation as an amplitude response curve (ARC), using the metric presented in

equation (5.15). In figure 5.4, we calculate the PRC and ARC for a state perturbation of three different strengths. The weakest perturbation results in type 1 (weak) resetting, in which the phase response curve is continuous, while the strongest perturbation results in type 0 (strong) resetting, in which the PRC is discontinuous. This transition occurs once the state perturbation is strong enough to push the trajectory over the unstable fixed point located at the middle of the limit cycle. For a perturbation of intermediate strength, at a critical phase the trajectory will be pushed close to the fixed point and take a long time to recover the steady state amplitude. The behavior at this singularity point is indicated by a sharp dip in the ARC for this perturbation strength.

Infinitesimal versions of response curves are often more general and easier to compute than those that track a specific perturbation strength [111]. We therefore derive an expression for the infinitesimal ARC, defined as

$$\frac{dA}{dx} := \lim_{\Delta x(0)\to 0} \frac{\Delta A\left(x(\tilde{t}),\ x^{\gamma}(\tilde{t}+\theta_0+\Delta\theta)\right)}{\Delta x(0)} = \int_0^{\infty} \lim_{\Delta x(0)\to 0} \frac{h(t)}{\Delta x(0)}\, dt. \tag{5.16}$$

While this quantity could be calculating by using a very small $\Delta x(0)$, it is more accurate and efficient to derive a direct method to calculate equation (5.16) using the ODE sensitivities defined in equation (1.5). For simplicity, we define $t_\theta = \tilde{t} + \theta_0$, which allows the time variable for the perturbation to vary from $0 \to \infty$ while tracking the appropriate phase on the limit cycle. Since $\Delta\theta \to 0$ as $\Delta x(0) \to 0$, we Taylor expand the limit cycle trajectory around t_θ:

$$x^{\gamma}(t_\theta + \Delta\theta) = x^{\gamma}(t_\theta) + \frac{dx^{\gamma}(t_\theta)}{d\theta}\Delta\theta + O(\Delta\theta^2) \tag{5.17}$$

$$= x^{\gamma}(t_\theta) + \tilde{f}\left(x^{\gamma}(t_\theta)\right)\Delta\theta + O(\Delta\theta^2). \tag{5.18}$$

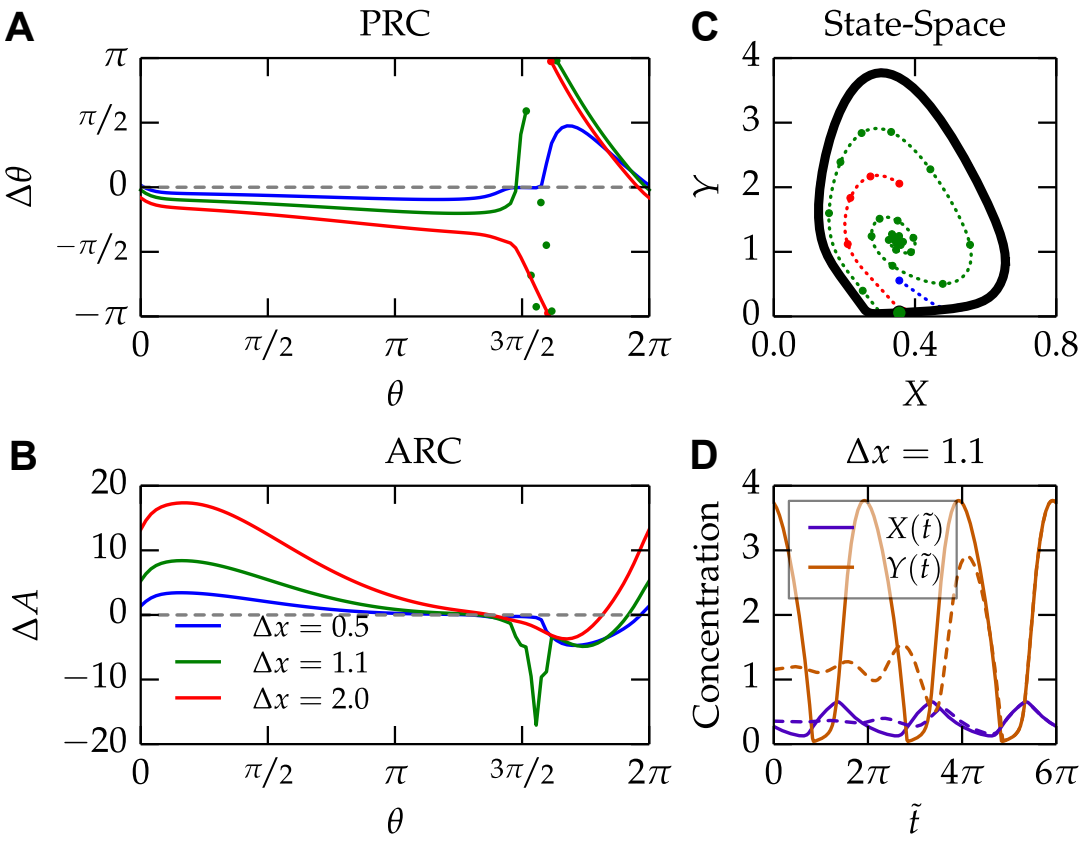


Figure 5.4: Phase and amplitude response curves describe how a single oscillator will respond to a perturbation. (**A-B**) The phase and amplitude response curves for three perturbations of increasing strengths demonstrate the transition from type 1 phase resetting for a weak stimulus (blue) to type 0 phase resetting from a strong stimulus (red). (**C**) The state space representation for the perturbations of varying strengths demonstrate how the trajectories return to the limit cycle. For the particular perturbation shown in green, the trajectory starts very close to the singularity, and therefore takes a long time to recover. (**D**) For an oscillator perturbed to the singularity, a strong amplitude reduction ensues. Here, the oscillator with an initial phase $\theta \approx 3\pi/2$ is perturbed by $\Delta x = 1.1$ at $\tilde{t} = 0$. Several cycles are required for normal amplitudes to be restored, corresponding to a dip in the ARC (C, green). Figure reprinted from [122].

Simplifying the integrand in equation (5.16),

$$\frac{h(t)}{\Delta x(0)} = \frac{1}{\Delta x(0)}\left[\left(x^{\gamma}(t_\theta) + \Delta x(\tilde{t}) - \mu\right)^2 - \left(x^{\gamma}(t_\theta + \Delta\theta) - \mu\right)^2\right] \tag{5.19}$$

$$= \frac{1}{\Delta x(0)}\left[\left(x^{\gamma}(t_\theta) + \Delta x(\tilde{t}) - \mu\right)^2 - \left(x^{\gamma}(t_\theta) + \tilde{f}\left(x^{\gamma}(t_\theta)\right)\Delta\theta - \mu\right)^2\right] \tag{5.20}$$

$$= \frac{1}{\Delta x(0)}\left[\left(\Delta x(\tilde{t}) - \tilde{f}\left(x^{\gamma}(t_\theta)\right)\Delta\theta\right)\left(\Delta x(\tilde{t}) + \tilde{f}\left(x^{\gamma}(t_\theta)\right)\Delta\theta + 2(x^{\gamma}(t_\theta) - \mu)\right)\right]. \tag{5.21}$$

Taking the limit of this integrand as $\Delta x(0) \to 0$ cancels several differential terms, and allows the remainder to be substituted with quantities that can be calculated using ODE sensitivity analysis:

$$\lim_{\Delta x(0)\to 0}\frac{h(t)}{\Delta x(0)} = 2\left(\lim_{\Delta x(0)\to 0}\frac{\Delta x(\tilde{t})}{\Delta x(0)} - \tilde{f}\left(x^{\gamma}(t_\theta)\right)\lim_{\Delta x(0)\to 0}\frac{\Delta\theta}{\Delta x(0)}\right)\left(x^{\gamma}(t_\theta) - \mu\right) \tag{5.22}$$

$$= 2\left(S(\tilde{t}) - \tilde{f}(x^{\gamma}(t_\theta))\frac{d\theta}{dx}\right)\left(x^{\gamma}(t_\theta) - \mu\right). \tag{5.23}$$

Note that in equation (5.23), $d\theta/dx$ represents the derivative of $\Delta\theta$ with respect to the perturbation, and is therefore a scalar quantity. The infinitesimal state-impulse ARC may therefore be calculated directly from the ODE sensitivity matrix and the PRC, allowing the amplitude change for an infinitesimal perturbation to be calculated exactly and efficiently:

$$\frac{dA}{dx} = \int_0^{\infty} 2\left(S(\tilde{t}) - \tilde{f}(x^{\gamma}(t_\theta))\frac{d\theta}{dx}\right)\left(x^{\gamma}(t_\theta) - \mu\right)\, dt. \tag{5.24}$$

Here, the first term of the integrand $(S - \tilde{f}\dot{\theta})$ tracks the distance from the perturbed trajectory to the limit cycle, which decays to zero as $t \to \infty$. The second term $(x^{\gamma} - \mu)$ weights this distance by whether or not the deviation occurs above or below the oscillatory mean, yielding negative amplitude changes when the trajectory is per-

turbed closer to the mean. Just as with the parameter-impulse PRC, the infinitesimal parameter-impulse ARC is defined as

$$\frac{d}{d\tilde{t}}\frac{dA}{dp} := \lim_{\tilde{d},\,\Delta p\to 0}\frac{\Delta A}{\tilde{d}\,\Delta p}, \tag{5.25}$$

and may be calculated from the state-impulse version with the following relationship:

$$\frac{d}{d\tilde{t}}\frac{dA}{dp} = \frac{dA}{dx}\frac{d\tilde{f}}{dp}. \tag{5.26}$$

As with equation (1.11), this equivalency reflects the fact that for a pulse of infinitely short duration, a parameter change is equivalent to changing the state of the system along the direction specified by the Jacobian. Convergence between the methods in equations (5.24-5.26) and finite-difference approaches is shown in figure 5.5, which demonstrates that the numerically-efficient differential ARCs remain representative even for moderate-strength perturbations.

5.3.3 Population-level response curves

We next briefly describe how to calculate the PRC and ARC for a phase-density model. These approaches are commonly used in understanding phase models [110, 124], and are presented here to match previous definitions for single cells. A phase transition curve, $g(\theta) = \theta + \Delta\theta$, maps the phase of an oscillator before perturbation to its phase after the perturbation. We denote the phase probability density function following a perturbation as $\hat{p}(\theta, \tilde{t})$, as shown in figure 5.6.

Since individual oscillators are neither created nor destroyed during the perturbation, it is possible - yet numerically difficult - to directly calculate the phase probability distribution following perturbation using the standard change of variables

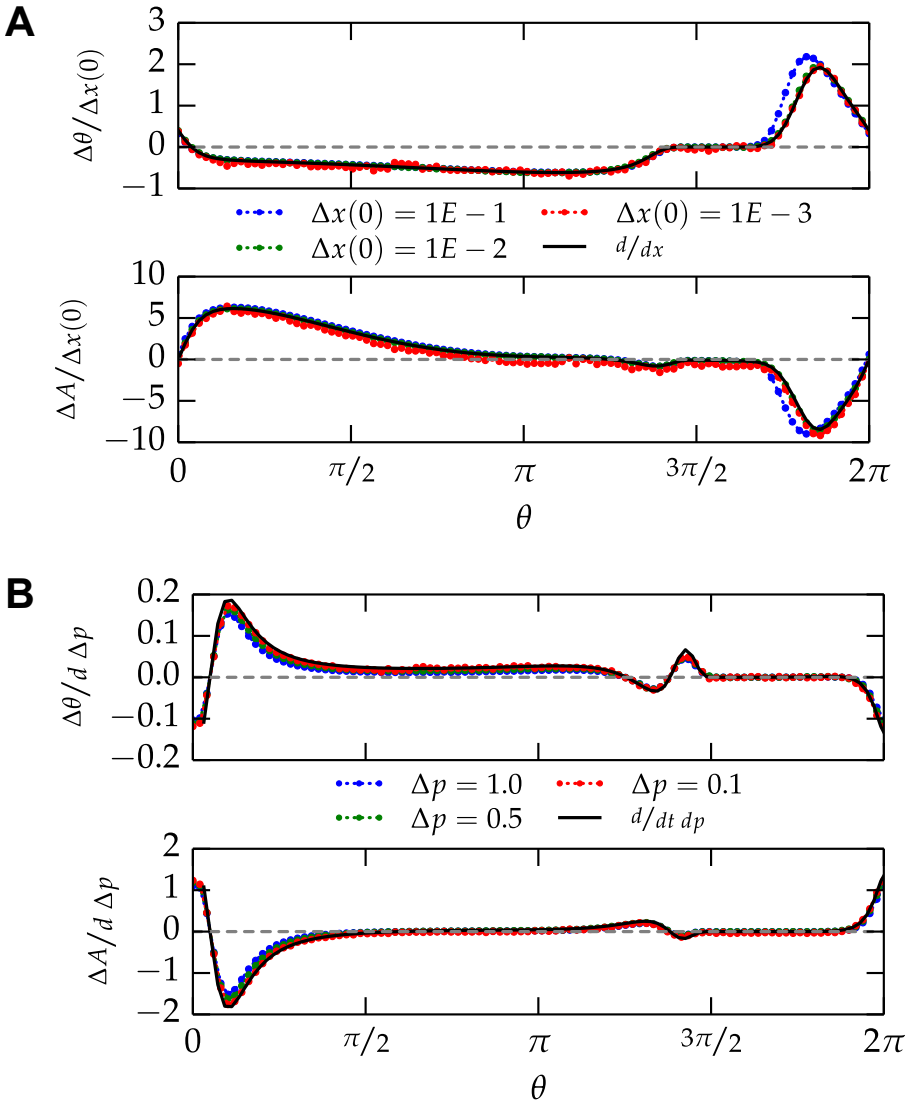


Figure 5.5: Convergence of finite-difference and differential methods. (**A**) Perturbations of decreasing strength to the state of the oscillator result in phase and amplitude response curves that match the differential limit. (**B**) Temporary perturbations of decreasing strength to a kinetic parameter. The duration of each perturbation is fixed to 0.2 radians. Finite difference approximations closely match the differential method, which is offset in phase by 0.1 radians to account for the nonzero pulse duration. Figure reprinted from [122].

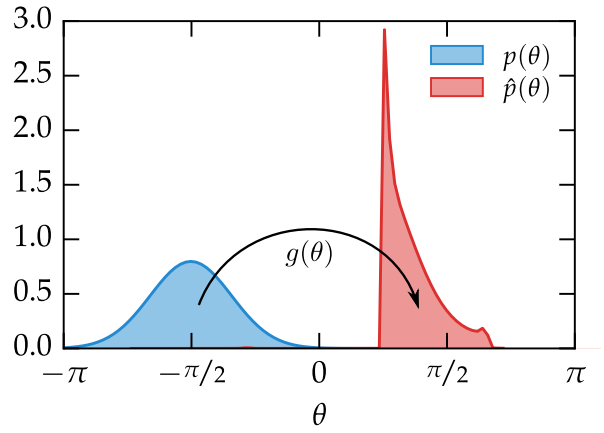


Figure 5.6: Perturbations change population-level phase synchrony. A perturbation giving rise to a phase transition curve $g(\theta)$ results in a changed phase probability density function, denoted as $\hat{p}(\theta, \tilde{t})$. Figure reprinted from [122].

relation:

$$\hat{p}(\theta, \tilde{t})\, dg(\theta) = p(\theta, \tilde{t})\, d\theta. \tag{5.27}$$

However, it is easier to estimate the mean and standard deviation of the perturbed population $\hat{p}(\theta, \tilde{t})$ using directional statistics [129]. A population defined on the unit circle can be described by a complex variable $z = \rho e^{i\bar{\theta}}$, where $\bar{\theta}$ is the mean phase and ρ, the synchronization index, is related to the standard deviation of the population. For $\rho = 1$, the population is clustered about one mean phase, while for $\rho = 0$ the population is evenly balanced across the unit circle. Complex variables for the population before and after perturbation can be calculated via

$$z := \int_0^{2\pi} e^{i\theta} p(\theta, \tilde{t})\, d\theta \tag{5.28}$$

$$\hat{z} := \int_0^{2\pi} e^{i\theta} \hat{p}(\theta, \tilde{t})\, d\theta = \int_0^{2\pi} e^{ig(\theta)} p(\theta, \tilde{t})\, d\theta. \tag{5.29}$$

In equation (5.29), we avoid calculating the perturbed population explicitly by instead integrating over the new phases at the prior population density function. Population-

level amplitude and phase responses can therefore be calculated by

$$\Delta\bar{\theta} = \angle z - \angle \hat{z} \tag{5.30}$$

$$\Delta\rho = |z| - |\hat{z}|. \tag{5.31}$$

Population-level PRCs and ARCs can be tabulated by solving equations (5.28-5.31) for populations $p(\theta, \tilde{t})$ with different mean phases. It is important to note that the ARC at the population level strongly depends on the slope of the PRC, as has been shown previously [110].

5.3.4 Population-level mean expression profiles

To efficiently capture the population-level effects of bioluminescence experiments, we couple the detailed single-cell ODE model to a phase-density model. Previous work has used limit cycle models to estimate population-level parameters, such as desynchronization rate, for phase-only models [127]. In this chapter, we use an ODE model to calculate the response to a perturbation at each phase, and subsequently take the weighted average of these responses according to the phases of the cells in the population.

Assuming each oscillator in the population follows the dynamics described by $x^{\gamma}(\theta)$, the unperturbed mean population-level expression, $\bar{x}(\tilde{t})$, can be found by taking the weighted average of the expression level over the current population:

$$\bar{x}(\tilde{t}) = \int_0^{2\pi} x^{\gamma}(\theta) p(\theta, \tilde{t})\, d\theta. \tag{5.32}$$

Phase-diffusion models can explain why the gradual damping from experimental population-level data closely resembles an exponentially damped sinusoid, a result

that has been shown experimentally [25] and computationally [127]. Our method similarly demonstrates exponential decay: for the idealized system $x^\gamma(\theta) = \cos(\theta)$ starting from a synchronized state:

$$\bar{x}(\tilde{t}) = \frac{1}{\sqrt{4\pi d\tilde{t}}} \int_{-\infty}^{\infty} \cos(x) \exp\left(-\frac{(x-\tilde{t})^2}{4d\tilde{t}}\right) \, dx \tag{5.33}$$

$$= \Re\left[\frac{1}{\sqrt{4\pi d\tilde{t}}} \int_{-\infty}^{\infty} \exp(ix) \exp\left(-\frac{(x-\tilde{t})^2}{4d\tilde{t}}\right) \, dx\right] \tag{5.34}$$

$$= \Re\left[e^{(i-d)\tilde{t}}\right] \tag{5.35}$$

$$= e^{-d\tilde{t}} \cos(\tilde{t}). \tag{5.36}$$

Due to the smoothing effect of phase diffusion, higher frequency sinusoidal components of the limit cycle are damped faster than lower frequency components, resulting in an exponentially decaying sinusoids even for limit cycles that are not sinusoidal in sufficiently disperse populations. The analytical result in equation (5.36) allows us to easily estimate the phase diffusivity parameter in equation (5.7) by simply fitting an exponentially damped sinusoid to detrended bioluminescence data.

Using our continuous approximation to population-level dynamics, we demonstrate an example unperturbed trajectory using a model adapted from [3], see model 6. In figure 5.7, an initially jagged population density smooths and widens over time as cells desynchronize, similar to the effects of diffusion. While each cell's expression level follows the limit cycle, the population amplitude gradually damps with time as cells with diverse phases are averaged together.

Next we describe how to calculate population-level mean expression following a perturbation. The perturbed trajectory $\hat{x}(\tilde{t})$ can be decomposed into contributions from two sources. First, for long times after the perturbation, each individual oscillator will have returned to the limit cycle $x^\gamma(\theta)$, but with a new population density

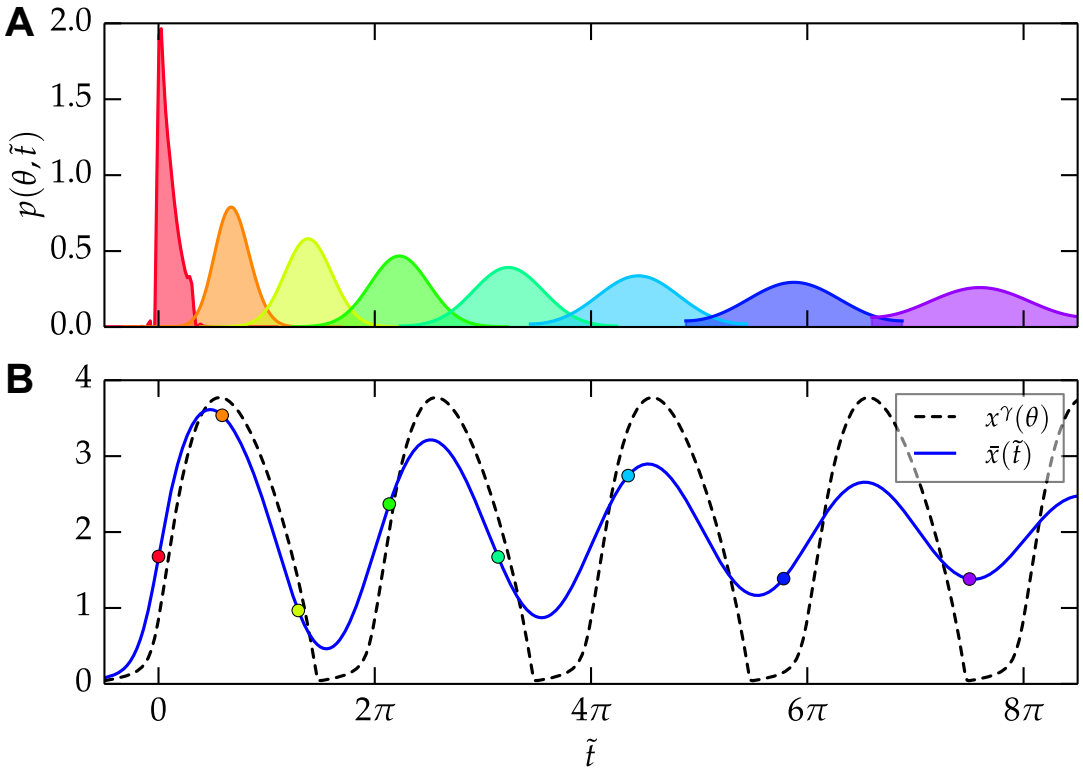


Figure 5.7: Synchrony affects population-level amplitude. (**A**) Stochastic fluctuations cause a population of cells to gradually desynchronize with time. The phase probability density, $p(\theta,\tilde{t})$, gradually widens as it advances in phase according to the mean period. (**B**) The mean amplitude from a population of oscillators is determined by the probability density function $p(\theta,\tilde{t})$ and the oscillator's limit cycle $x^{\gamma}(\theta)$. As time passes, population-level rhythms resemble an exponentially damped sinusoid. Figure reprinted from [122].

Model 6: Model adapted from Novák & Tyson, 2008 [3]. Used for ARC demonstrations in figures 5.3-5.7 ($P = 4$) and figure 5.8 ($P = 2$). It should be noted that the use of the stochastic simulation algorithm for reactions with non-elementary propensities may inaccurately represent the noise of the full system. In this case, we simply fit the noise characteristics of the approximated model (using the volume parameter) to yield physiologically realistic desynchronization rates. The volume of the stochastic simulations was chosen as $\Omega = 250$. Figure reprinted from [122].

$$\frac{dX}{dt} = \frac{1}{1+Y} - X$$
$$\frac{dY}{dt} = k_t\, X - k_d\, Y - \frac{Y}{\alpha_0 + \alpha_1\, Y + \alpha_2 Y^2}$$

Parameter	Value
k_t	20
k_d	1
P	4 (or 2)

Parameter	Value
α_0	0.005
α_1	0.05
α_2	0.1

$\hat{p}(\theta, \tilde{t})$. This steady-state perturbed trajectory $\hat{x}_{ss}(\tilde{t})$ can be found by

$$\hat{x}_{ss}(\tilde{t}) = \int_0^{2\pi} x^{\gamma}(\theta)\hat{p}(\theta, \tilde{t})\, d\theta. \tag{5.37}$$

For very long times following perturbation, the new phase probability density could be approximated using the initial mean and standard deviation found through equation (5.29), as jagged profiles in the population density will eventually smooth to a normal distribution. However, for shorter times following perturbation, $\hat{p}(\theta, \tilde{t})$ must be calculated numerically.

The second contribution to the perturbed population trajectory comes from deviations from limit cycle oscillations in each cell. We calculate the population-level effect of these deviations by averaging over the deviations that occur at each phase. We define the deviation trajectory $\delta x(\theta, \tilde{t})$ for each phase as the distance between the

perturbed trajectory and the phase-adjusted reference:

$$\delta x(\theta_0, \tilde{t}) := x(\tilde{t}) - x^{\gamma}(\tilde{t} + \theta_0 + \Delta\theta) \tag{5.38}$$

$$\therefore \lim_{\tilde{t}\to\infty} \delta x(\theta, \tilde{t}) = 0. \tag{5.39}$$

Since the perturbed trajectory ultimately converges with the phase-adjusted reference, deviations will converge to zero. Since the phase change, $\Delta\theta$, associated with a perturbation at each phase is likely not known prior to calculating the perturbed trajectory, it is difficult to tabulate deviation trajectories associated with each final phase, $\theta_0 + \Delta\theta$. It is therefore more straightforward to find the average effect of single-cell perturbations at the population level by weighting the deviations by the phase density function prior to perturbation. The population-level response to a perturbation, $\hat{x}(\tilde{t})$, is therefore defined as

$$\hat{x}(\tilde{t}) = \int_0^{2\pi} x^{\gamma}(\theta)\hat{p}(\theta, \tilde{t}) + \delta x(\theta, \tilde{t})p(\theta, \tilde{t})\, d\theta. \tag{5.40}$$

Here the first term is equivalent to $\hat{x}_{ss}(\tilde{t})$, the steady-state perturbed trajectory, while the second term decays to zero as individual oscillators return to their steady-state amplitude. The accuracy of the continuous phase-diffusion model was tested by explicitly simulating a perturbation to a population of 225 uncoupled oscillators using a stochastic simulation algorithm [38, 130]. Good agreement between the stochastic population and continuous approximation was shown in phase shift, single-cell level amplitude change, and alteration of population synchrony (figure 5.8). These results indicate the use of a continuous probability function is justified in the case of cultured cellular reporter systems, which may contain up to 10^5 individual cells per culture [25]. In addition, the method was verified by similar simulations us-

ing the Oregonator (model 7, [4]), a model containing only mass-action terms, and a detailed mechanistic model of circadian rhythms (model 1) at a variety of phase timings (figure 5.9 and figure 5.10). Good agreement is seen between the continuous and stochastic simulations, indicating the proposed method is suitable for predicting the population-level responses for a variety of model types. The slightly reduced accuracy seen in the larger model 1 demonstrates a limitation of the method, which should be considered before applying the method to extreme cases. As equation (5.39) tabulates the response of the system from the limit cycle, systems which are perturbed from states far from the deterministic limit cycle will not be captured appropriately. Therefore, the difference in accuracy between model 1 and model 7 likely comes from increased deviation about the deterministic limit cycle in a greater number of spatial dimensions. The continuous approximation is therefore able to capture the population-level dynamics of a wide variety of limit cycle models and perturbations.

We next demonstrate the usefulness of single-cell and population level ARCs in predicting the mean response of a population. Using a detailed single-cell model (model 1), we calculate the response of the system, specifically the average CRY2 protein expression level, to a temporary increase in *Per* transcription rate. In part A of figure 5.11, we plot the resulting single-cell and population-level response curves. In this case, the population-level phase change is a slightly smoothed version of the single-cell PRC, since the population has an averaging effect on incoming perturbations with each cell receiving the input at a slightly different internal phase. This smoothing follows from the consistent definition of phase between the single-cell and population level. In the ARC plot, however, the shape of the single-cell and population-based ARC are different, since they describe different types of changes (finite vs. sustained) in the output trajectories.

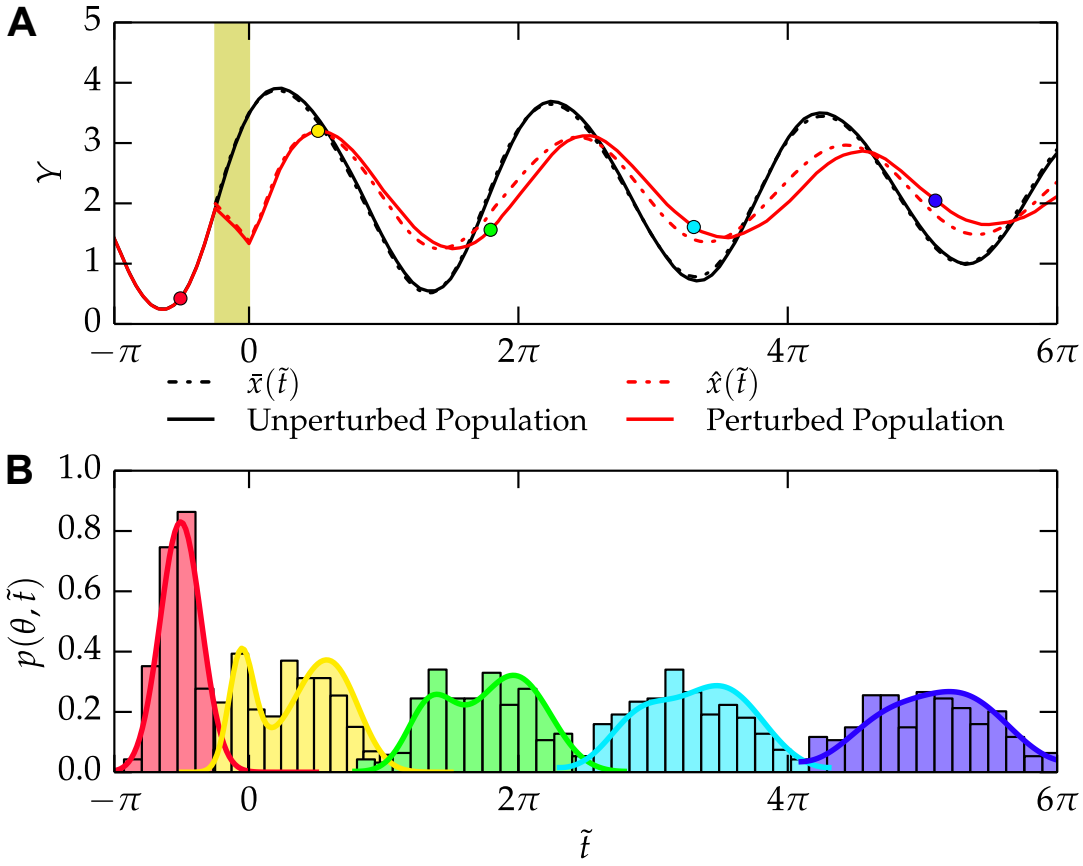


Figure 5.8: Approximation of an explicit stochastic population by continuous methods. (**A**) A population of 225 stochastic oscillators was simulated using the Gillespie stochastic simulation algorithm (SSA). The mean protein expression, Y, is plotted as a function of time. A 50% reduction in the protein translation rate to each of the oscillators is applied from $\hat{t} = -\pi/4 \rightarrow 0$, resulting in both single-cell and population-level amplitude change (red solid line). This population is approximated by the continuous methods described in this manuscript, in which the decay parameter $d = 0.025$ is estimated to match the stochastic-induced desynchrony of the control population (black solid line). The initial standard deviation $\sigma_0 = 0.48$ is similarly matched to the stochastic population. The resulting predicted unperturbed and perturbed trajectories, $\bar{x}(\hat{t})$ and $\hat{x}(\hat{t})$ respectively, closely match the stochastically modeled values. (**B**) Phase histograms for the stochastic population are shown at several phases, both before and after the desynchronizing perturbation. Phase probability-density functions for the continuous approximation are also shown, with close agreement between stochastic and continuous simulations. This close approximation validates the use of ODE models and phase-diffusion populations in deriving amplitude and phase-response behavior for networks of uncoupled cells. Figure reprinted from [122].

Model 7: The Oregonator model [4]. The parameters used are as presented in [38]. Used in figure 5.9 as an example of a limit cycle oscillator using only mass-action kinetics. Figure reprinted from [122].

The individual chemical reactions (used for stochastic simulation) are:

$$
\begin{aligned}
X_1 + Y_2 &\xrightarrow{c_1} Y_1 \\
Y_1 + Y_2 &\xrightarrow{c_2} Z_1 \\
X_2 + Y_1 &\xrightarrow{c_3} 2Y_1 + Y_3 \\
2Y_1 &\xrightarrow{c_4} Z_2 \\
X_3 + Y_3 &\xrightarrow{c_5} Y_2
\end{aligned}
$$

These reactions can be converted to a standard ODE model, in which only the intermediate variables Y are considered, as:

$$
\begin{aligned}
\frac{d\,Y1}{dt} &= c1x1 \cdot Y2 - c2 \cdot Y1 \cdot Y2 + c3x2 \cdot Y1 - c4 \cdot Y1^2 \\
\frac{d\,Y2}{dt} &= -c1x1 \cdot Y2 - c2 \cdot Y1 \cdot Y2 + c5x3 \cdot Y3 \\
\frac{d\,Y3}{dt} &= c3x2 \cdot Y1 - c5x3 \cdot Y3
\end{aligned}
$$

Parameter	Value
$c1x1$	2
$c2$	0.1
$c3x2$	104
$c4$	0.016
$c5x3$	26

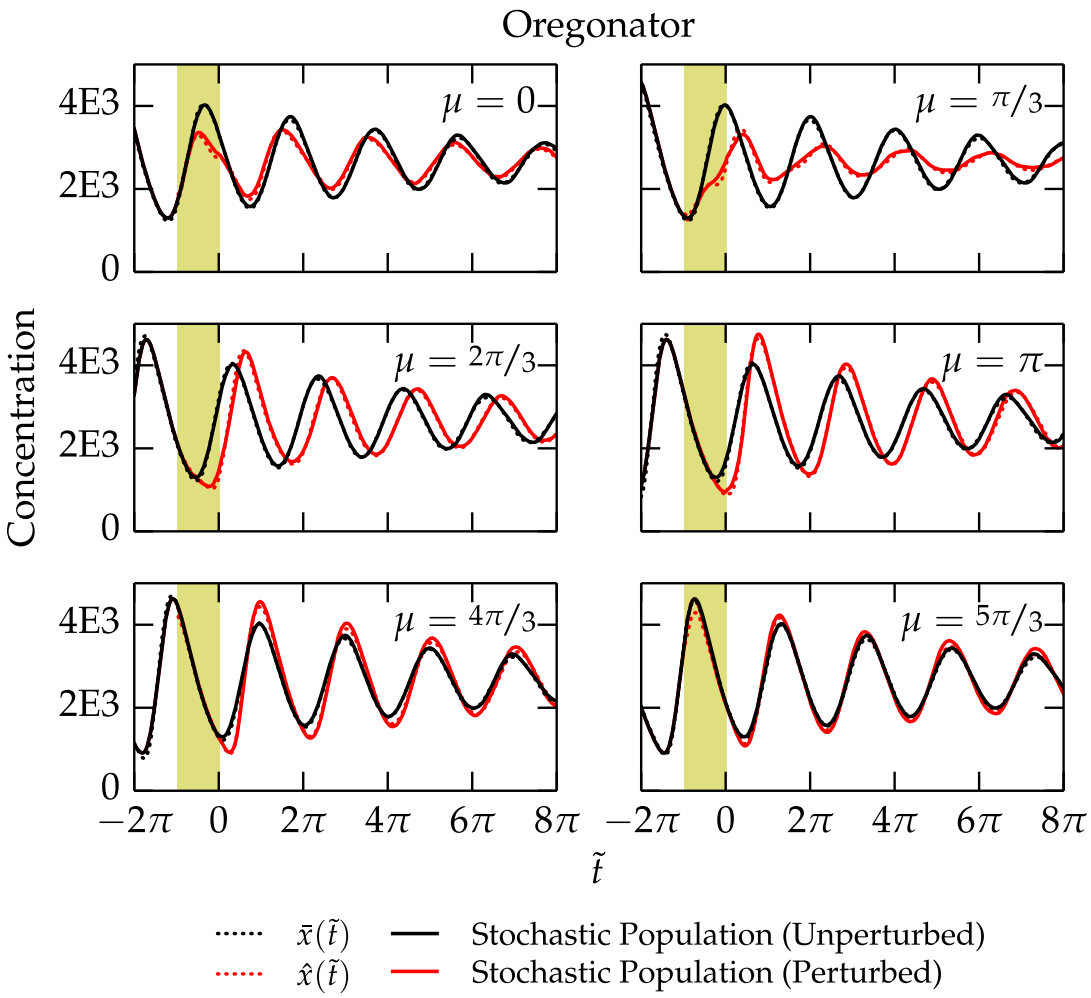


Figure 5.9: Validation of continuous methods with the Oregonator model. The Oregonator model (model 7) was tested as an example of a mass-action limit cycle oscillator. A population of 2000 oscillators was perturbed at size different mean phases, μ, by increasing the $c2$ parameter by 50% for $d = \pi$ (highlighted region). Only the average $Y2$ variable is plotted. Figure reprinted from [122].

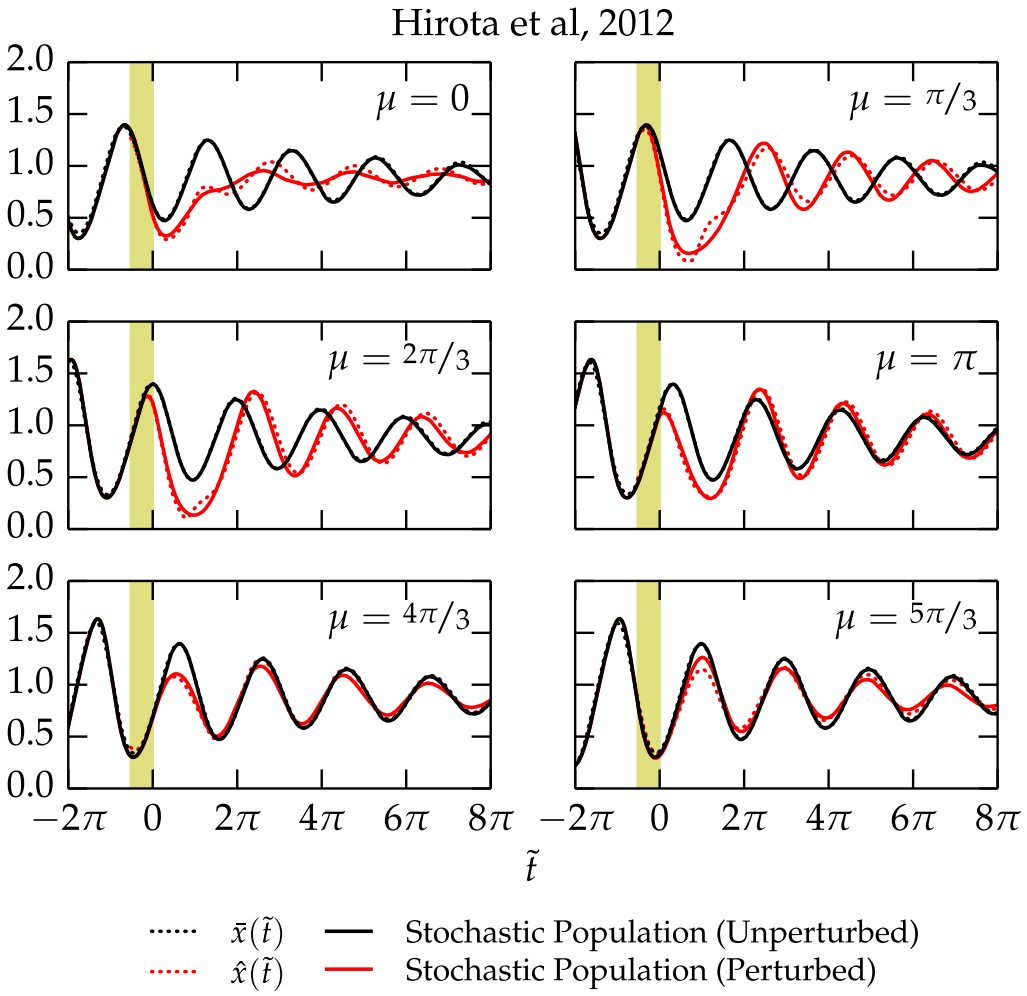


Figure 5.10: Validation of continuous methods with the model from Hirota *et al.*, 2012. The model described in [49] (model 1) was similarly tested at a variety of mean phases by reducing the *vdp* parameter by 28.5% and plotting the resulting *c2* state variable (as in figure 5.12). Parameters for the continuous approximations were found by estimating the phase, initial standard deviation, period, and phase diffusivity of the control population. Figure reprinted from [122].

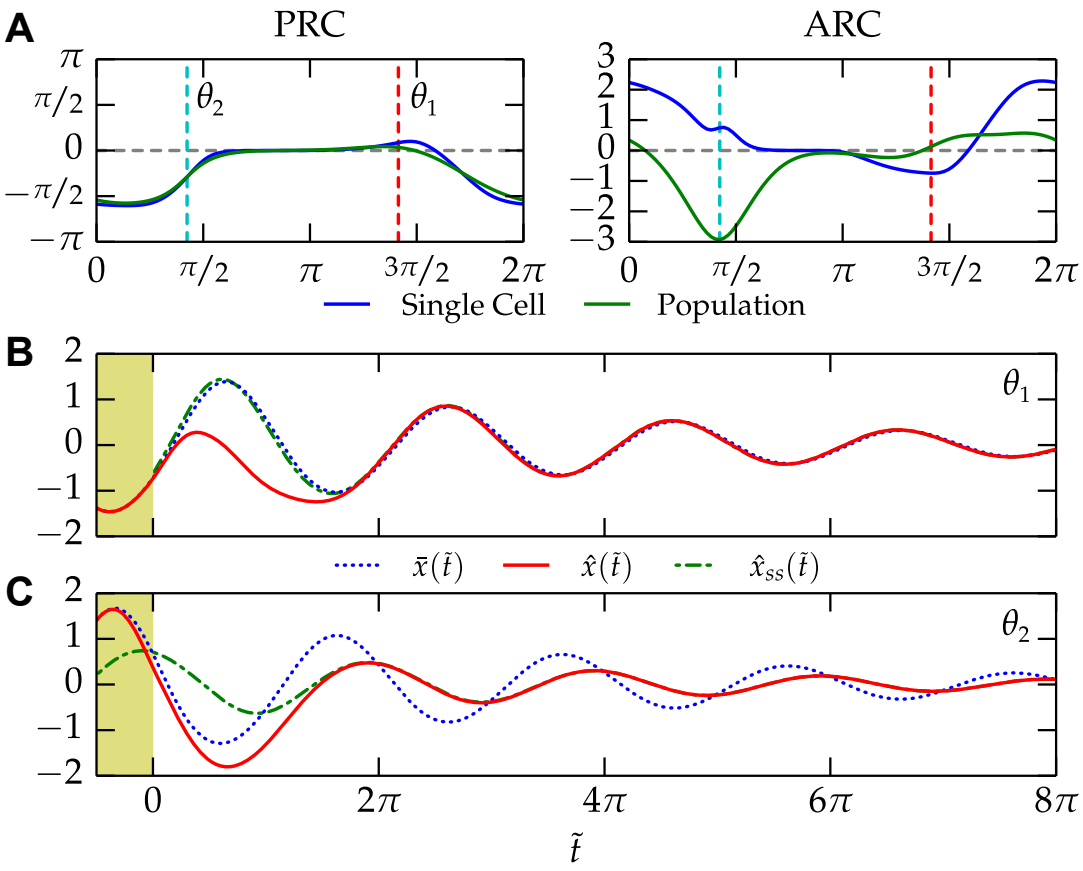


Figure 5.11: Response curves describe different types of amplitude change. A population of cells is simulated using a detailed model of circadian rhythms (model 1). Each cell in the population is subjected to a 20% increase in the *vtp* parameter (*Per* transcription rate) for $d = \pi/2$, inducing both a phase and amplitude change. (**A**) The phase response curves are shown for both the single limit cycle oscillator and population average, left. The population-level ARC is shown together with the single-cell amplitude response, with both curves normalized to $\sigma = 1$. (**B-C**) Population-level rhythms, $\hat{x}(\tilde{t})$ resulting from perturbations given at the indicated phases in A (red and cyan lines, respectively). The population transitions from the unperturbed population, $\bar{x}(\tilde{t})$, to the steady-state perturbed population, $\hat{x}_{ss}(\tilde{t})$. (**B**) A perturbation at this phase yields a transient amplitude reduction resulting from perturbations at the single-cell level, however, there is little change in population synchrony. Population-level rhythms are therefore transiently damped before regaining normal amplitudes. (**C**) Oscillators are desynchronized, but with a transient increase in limit cycle amplitude. It therefore takes some time before the amplitude reduction from desynchrony is realized. Figure reprinted from [122].

Using these response curves, we demonstrate characteristic features of amplitude change mediated at the single-cell and population-level. In part B of figure 5.11, the perturbation at θ_1 does not cause a significant change in phase or population synchrony, yet strongly reduces single-cell amplitudes. The resulting population-level trajectory appears damped for the first cycle, until the oscillators within the population return to the limit cycle and normal amplitudes are restored. In part C, the perturbation at θ_2 reduces population synchrony while increasing single-cell amplitudes, yielding amplitudes that are transiently increased before the population settles to a lower-amplitude, desynchronized trajectory.

These examples demonstrate the qualitative differences between amplitude change mediated at the single-cell and population level. Their characteristic features allow us to distinguish between these two sources of amplitude change without explicitly recording single-cell amplitudes: amplitude change induced at the population-level will be sustained, but may be masked in the short term by changes to single-cell level amplitudes. Similarly, a change induced at the single-cell level will be evident only at short time-scales. These results underscore the importance in considering both single-cell and population-mediated amplitude change in predicting the effects of daily stimuli on clock amplitudes.

5.3.5 Application to Experimental Data

We conclude by using the proposed modeling framework to reconcile two experimental studies that measured the effect of light on the circadian amplitude of photosensitive fibroblast cells [109, 110]. Both studies sought to determine the main factor of amplitude reduction following a transient light pulse, in which either desynchrony alone [110] or a combination of single-cell amplitude reduction and desynchrony [109] was identified as the dominant factor.

Data on the response of melanopsin-responsive and control NIH3T3 mouse fibroblasts cells is shown in figure 5.12 [110]. The experimental data is de-noised using a discrete wavelet transform [131]. In the experiment, a light pulse is given to desynchronize a colony of cells, resulting in drastically reduced amplitudes. A second (longer) light pulse subsequently re-synchronizes the population, resulting in increased amplitudes. To capture this phenomenon *in silico*, we used a recent model of the core circadian feedback circuit [49]. In order to find parameters for the phase density function, we assumed an initial phase population with $\sigma = 0.1$ and calculated a phase diffusivity parameter $d = 0.104$ from the experimental control trajectory.

The next step in capturing the experimental data was finding a perturbation capable of desynchronizing the system. To find such a perturbation, we calculated population-level ARCs at several different knockdown strengths for each parameter in the model. We selected the degradation rate of *Per* mRNA from the feasible parameters due to PER's known induction by CREB following photo-perturbation [132]. While not an exact match of the experimental system, reducing degradation rate yields a similar response to increasing transcription. The parameter was reduced by 28.5% during the light pulses, a value fit to maximize light-induced desynchrony. The output variable, PER2-luciferase in the experimental system, was chosen to be represented by *Cry1* mRNA - an E box activated gene that is buffered from the direct effects of the parameter perturbation. The single-cell and population-level response curves for this parameter choice are shown in figure 5.12, which demonstrates the strong desynchronization that occurs at $\theta \approx \pi/4$. Since the reporter used in the experimental system likely has a phase lag from the corresponding mRNA, we did not try to match the initial phase of the simulation to experiment. Instead, the initial phase of simulation was chosen such that the first pulse occurred when the system was at the phase that corresponded with the minimum of the population-level ARC.

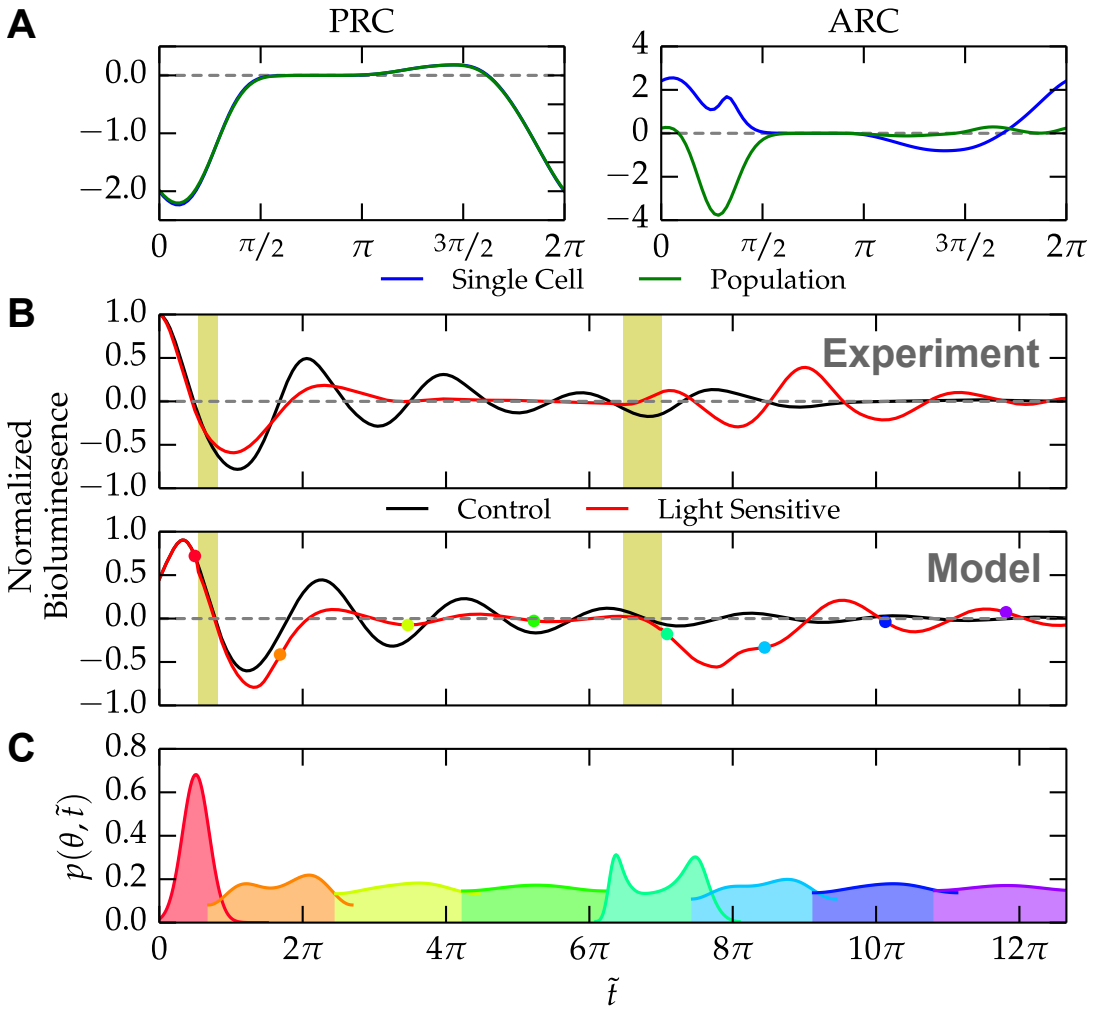


Figure 5.12: Method allows direct comparison between model results and experimental bioluminescence profiles. (**A**) Response curves for the parameter perturbation used to model the first light pulse in (B). The initial phase of the system is chosen to coincide with the strong dip in the population-level ARC (left). (**B**) De-noised bioluminescence data from Ukai *et al.* 2007 (top) demonstrates population-level amplitude change resulting from two light pulses [110]. Circadian oscillations are suppressed by the first light pulse and rescued by the second. These perturbations were reproduced *in silico* (bottom) using the model from Hirota *et al.* 2012. The experimental results are qualitatively captured by the model, demonstrating the suitability of the method in capturing bioluminescence experiments. The relative amplitude of the model equations has been scaled to match that of the normalized bioluminescence profiles. (**C**) Example probability density functions for the model trajectory shown in (A). The initial population (red) is effectively desynchronized (orange) following the first light pulse. Figure reprinted from [122].

Population-level trajectories were found by solving for the resulting limit cycle trajectories at every phase, and subsequently finding the weighted average of these trajectories according the phase density (as described in equations (5.32 and 5.40)). The simulated control and perturbed trajectories, Fig. 5B, closely match the experimental results, in which the model captures both the phase shifts and amplitude modulation of the light pulses. The phase probability density function for the light-sensitive model trajectory is shown in Fig. 5C for several representative time points. Changes to the synchronization of the population from each perturbation are readily apparent, with both perturbations inducing a bimodal distribution in the phase density. While experimental data on individual cell phases was unavailable for this data set, bimodal distributions in SCN neuron firing following a phase shift have previously been seen experimentally [133]. As higher-frequency features, these bimodal distributions dissipate as the phases diffuse. Amplitude change in this case is mediated both at the population and single-cell level, as indicated by their representative ARCs. The contributions from each source are summarized in Fig. S4.

Several quantitative differences between the model and experiment do appear. Most importantly, the acute amplitude change following the light pulse is not correctly captured by the single-cell model. From the experimental system, it appears the light pulse temporarily reduces PER2-luciferase amplitudes immediately following perturbation, separate from the overall change in population synchrony. This result is most obvious after the second light pulse, where rhythms require some time before their maximum amplitude is reached. Such a delay is indicative of perturbations to the limit cycle system, suggesting a contribution from single-cells in determining the population amplitude. Light-induced amplitude suppression in fibroblasts is therefore likely mediated at the single-cell level at short time-scales, and at the population level for longer time-scales. To improve the fit of the model to

the data, the response curves could be fit to match experimentally predicted values. Infinitesimal PRCs and ARCs calculated at the single-cell level are numerically efficient to evaluate, and could be more readily incorporated into a parameter estimation algorithm than explicit stochastic simulations of a population of oscillators.

5.4 Discussion

In this chapter, we have described new tools for understanding amplitude change in independent cells based on analytically tractable methods for simulating a large population of oscillators. These tools synthesize single-cell level ODE sensitivity analysis metrics with population-level measures of synchrony, allowing the quantification of clock amplitude at multiple levels of biological organization. Using these tools, we have demonstrated how ODE models can be directly compared to experimental bioluminescence profiles, which currently serve as a standard system to investigate clock perturbations. More generally, we have demonstrated the differences between acute and prolonged changes in amplitude following temporary perturbations, and characterized the mechanisms that give rise to each.

5.4.1 Application to pharmacological control

While our examples have focused on light-mediated perturbations, these approaches could also be applied to pharmacological perturbations. Some difficulty arises in obtaining such data in cultured cells, however, as entraining perturbations would require that pharmacological agents be introduced for only a finite duration. As medium or temperature changes are often enough to resynchronize cultured cells, a transient application is difficult to achieve experimentally. The search for clock-enhancing molecules has therefore tended to focus on constant drug concentrations:

for instance, dose-dependent period or amplitude change following inhibition of CK1δ or similar targets [48]. For *in vivo* systems, however, pharmacokinetics dictates a finite duration of action for both naturally secreted hormones and pharmacological therapies. More effective treatments might therefore be designed by explicitly accounting for such a transient response, perturbing peripheral clocks at the right phase to induce resynchronization and an increase in single-cell level amplitudes.

Maintaining robust rhythmicity in peripheral oscillators has been linked to protection against metabolic disease. It is therefore interesting that liver cells have not developed a stronger mechanism of intercellular coupling, such as in the SCN, to maintain robust amplitudes in the absence of external cues. Perhaps the ability to quickly re-entrain to a phase shift, which is often slowed by coupling [105], has historically been more advantageous than protection against highly variable food intake. Regardless, the necessities of modern society often dictate irregular circadian-metabolic regimes that damp peripheral clock oscillations. For such instances, we hope that the mathematical theory presented here will permit the development of optimized treatment regimes.

5.4.2 Damping rate is correlated with single-cell noise

One result from this chapter, which has previously been discussed in the circadian literature [127], is that in noninteracting populations, stochastic noise at the single-cell level affects the damping rate of a population of oscillators. Stochastic noise is often a relevant circadian parameter, as it describes the precision of the underlying oscillators. The following chapter discusses how population-level recordings, even those from high-throughput circadian screens, can be used to gain estimates of stochastic noise at the single-cell level. Using such an approach, it is possible to gain genome-wide insight into how noise is regulated in the cell-autonomous feedback

loop.

Chapter 6

Quantifying stochastic noise in cultured circadian reporter cells[1]

6.1 Background

In the previous chapter, mathematical definitions governing the response of a population of oscillators to a temporary perturbation were discussed. Temporary perturbations, arising perhaps from a single dose of drug that is quickly cleared from the body, are promising routes for increasing amplitude in peripheral tissues. However, because of the experimental difficulty associated with studying such perturbations in a high-throughput manner, most research has focused on studying the effects of constant concentrations of drugs or gene knockdowns. This chapter takes a closer look at the insight into clock kinetics that might be gained from such existing studies, specifically with regard to transcriptional noise.

[1]Portions of this chapter have been submitted for publication as P. C. St. John and F. J. Doyle III, "Quantifying stochastic noise in cultured circadian reporter cells"

6.1.1 Motivation

As most organisms exist in an environment that changes predictably with a 24-hour period, highly optimized genetic circuits turn on and off the production of key regulatory proteins to anticipate the day/night cycle. In humans, the demands of a modern society have required that we deviate from this evolutionarily prescribed sleep and feeding schedule, resulting in increased long-term risks of metabolic disease. There is therefore a desire to find pharmacological treatments that would restore the normal functioning of our circadian clock despite irregular behavioral schedules. One aspect of these treatments that is often overlooked in searching for candidate drugs is how these treatments might affect the accuracy of the circadian timing system. Since each cell in the body keeps its own internal time, increasing the sloppiness of the clock machinery can result in tissues not being able to agree on what time to turn on protein production. Recording the time of each cell is possible but difficult; as a result single-cell approaches cannot be scaled up to high-throughput searches. In this chapter, we show that it is possible to estimate how much the noise of a system has changed by looking only at the averaged protein production of an entire population of cells. Such an approach allows us to analyze prior data for high-throughput screens, and show that the unperturbed clock has been highly optimized to be both accurate and high amplitude.

6.1.2 Transcriptional noise

A detailed understanding of the underlying transcriptional mechanisms is essential for the development of circadian therapeutics to be successful. The functional roles of different genes in circadian regulation have traditionally been studied using behavioral-level data and genetic knockout experiments [20]. Bioluminescence-

based cellular circadian reporters offer a more direct view of the gene regulatory network [134] and are amenable to high-throughput screens, allowing genome-wide exploration into factors that affect circadian rhythmicity [22]. Additionally, cultured circadian reporter cells allow the change in transcriptional amplitude following a perturbation to be quantified. This additional parameter has proven useful in differentiating between perturbations with the same effect on period [96] and has aided the search for small-molecule therapeutics to boost clock amplitude [48].

Bioluminescence rhythms at the cell culture or tissue-level are the result of the collective behavior of thousands of cells. Transcription at the single-cell level is strongly affected by the low molecular counts of the mRNA and protein species involved. As a result, bioluminescence traces of individual cells are stochastic, with significant variability in both amplitude and period length [25]. It is currently thought that coupling between circadian oscillations outside the SCN, such as in peripheral tissues *in vivo* or cultured reporter cells *in vitro*, is very weak or absent entirely [24, 135]. Additionally, it has been shown that in cell cultures that lack cell-to-cell coupling, stochastic noise at the single-cell level is manifested in damped oscillations at the population-level as individual oscillators gradually drift out of phase [25, 136].

Despite the averaging that occurs at the population-level, cell-autonomous stochastic noise plays an important role in determining the function of the circadian oscillator. A recent study showed that stochasticity is critical to the population-level response to a neuropeptide and forms the basis for how the SCN entrains to light-mediated cues [121]. Additional studies have argued that the basis of single-cell rhythmicity depends on stochastic noise, as models of deterministically damped oscillators, when simulated stochastically, better capture the noise characteristics seen in single-cell fibroblast data [137]. Despite the importance of single-cell stochasticity in circadian rhythms, measuring stochastic noise currently requires careful prepara-

tion, recording, and image processing of individual cells [138]. As a result, while circadian perturbations have been postulated to affect single-cell stochasticity [127], no study has experimentally quantified changes to single-cell noise as a result of a small molecule or genetic perturbation.

6.1.3 Required assumptions on single-cell dynamics

In this chapter, we demonstrate that changes to single-cell stochasticity can be reliably inferred from the changes to the damping rate of population-level bioluminescence recordings of cultured circadian reporters. Our method assumes that oscillations in individual cells are both

- independent (no significant cell-to-cell coupling) [24, 135],
- and sustained (do not damp on a single-cell basis) [25, 136],

which have been shown to hold experimentally for cultured fibroblast cells. We demonstrate the validity and usefulness of such an approach on several types of circadian data. First, we show using single-cell fibroblast data that stochastic noise is directly related to population-level damping. Next, we show that a small-molecule modulator is able to change damping rate in a dose-dependent fashion, and verify using a mathematical model that changes to single-cell stochasticity is a likely mechanism. Finally, we calculate the genome-wide effects of siRNA knockdown on single-cell noise, and demonstrate that population-level damping rate is independent of other circadian parameters, such as period or amplitude. Using this additional information, we show that circadian rhythms have likely evolved to an optimal point of high amplitude and low stochastic noise. Our results should prove especially important in the future search for small molecule circadian therapeutics, as it allows the

effect of candidate drugs on single-cell noise to be quantified in a high-throughput manner.

6.2 Methods

6.2.1 Fitting a damped sinusoid to experimental data

Raw experimental data $x(t_i),\ i \in \{0, \ldots, N-1\}$ are first detrended using Hodrick-Prescott filter with a smoothing parameter $\gamma = 0.05\,(^{24\,\mathrm{hrs}}/_{T_s})^4$, in which T_s is the sampling rate (in hours) [139]. The detrended data are then filtered using a low-pass filter to remove high-frequency noise (forward-backward Butterworth filter with $n = 5,\ w_c = 0.1$). We denote the detrended and filtered experimental data by $y(t_i)$. A damped sinusoid, specified by:

$$\hat{y}(t) = Ae^{-dt}\sin\left(\frac{2\pi t}{T} + \theta\right)$$

is then fit to the filtered data. For numerical efficiency, the period, T, and damping rate, d, parameters are fit first using a matrix pencil method [140], reviewed in [141]. Amplitude, A, and phase, θ, parameters are subsequently fit using a linear least-squares regression. Overall R^2 values for the regression were calculated from the residual error between the detrended data and fitted sinusoid:

$$R^2 = 1 - \frac{\sum_{i=0}^{N-1}(y(t_i) - \hat{y}(t_i))^2}{\sum_{i=0}^{N-1}(y(t_i) - \bar{y}(t_i))^2}$$

in which $\bar{y}(t_i)$ represents the mean of the detrended data.

6.2.2 Processing single-cell bioluminescence data

Single-cell bioluminescence data for 79 cells was obtained from Leise *et al.*, 2012 [138]. As was done in the original study, a discrete wavelet transform (using PyWavelets, `http://www.pybytes.com/pywavelets`) was performed to detrend and remove noise.

Sorting cells by noise level

As in the original study, various parameters describing the average noise level of each cell were collected. Traces were denoised and detrended by keeping only the $(8\text{hr}, 258\text{hr})$ wavelet components. From these smoothed trajectories, a Hilbert transform was used estimate points at which the phase crossed 0 to find period and amplitude coefficients of variation. An additional noise parameter, the standard deviation in the $(1\text{hr}, 8\text{hr})$ wavelet components divided by the overall rhythm amplitude, was used to quantify the high-frequency noise of the system. From these three noise variables, a combined noise metric was constructed by projecting the variables along their first principle component (using scikit-learn, `http://scikit-learn.org/`). Cells were ranked according to this combined noise metric, and a high-noise group and low-noise group were constructed by taking the 39 highest-noise and lowest-noise cells, respectively. The raw bioluminescence profiles were not initially synchronized, so the traces were offset to have the same starting phase in order to simulate the gradual desynchronization of a group of oscillators. This was accomplished by starting each trace at the first phase zero-crossing, found using a Hilbert transform.

Bootstrap estimations of the damping rate difference

Averaged traces for low and high-noise group displayed a damped sinusoidal rhythm. The first 4 days of rhythms ($T_s = 0.5,\ N = 192$) were fit using a damped

sinusoid. To ensure the difference in damping rate between groups was statistically significant, a bootstrap analysis was performed. In each of $10,000$ bootstrap trials, cells were randomly assigned evenly to either the low-noise or the high-noise group (one cell was randomly omitted in each trial to ensure even group sizes). The absolute difference in damping rate between the two populations was recorded to yield a two-tailed test. The observed test statistic, $|d_h - d_l| = 6.65 \times 10^{-3}$, was found to be significant at the $\alpha = 0.05$ confidence threshold ($p = 0.0264$).

6.2.3 Quantifying dose-dependent effects of small molecule modulators

Bioluminescence traces ($T_s = 1.67,\ N = 71$) with increasing small molecule concentrations were fit with a damped sinusoid using the method described in a previous section. Because the small molecules were toxic at very high concentrations, experiments were removed from further analysis where $R^2 < 0.80$ (figure 6.1).

6.2.4 *In silico* prediction of small molecule experiments

A previously published mathematical model of circadian rhythms [49] was used to predict the effects on population damping rate from the dose-dependent small molecule experiments. The parameters used to capture the effects of each small molecule were the same as described previously [96]. The model was converted to a stochastic biochemical system and subsequently simulated using StochKit2 [130] (via GillesPy, `https://github.com/JohnAbel/gillespy`). Population-level rhythms were found by taking the average of $1,000$ noninteracting oscillators, starting at identical initial conditions. The only parameter left unspecified by the deterministic model was the cell volume, Ω, which controlled the amount of noise in the system. For

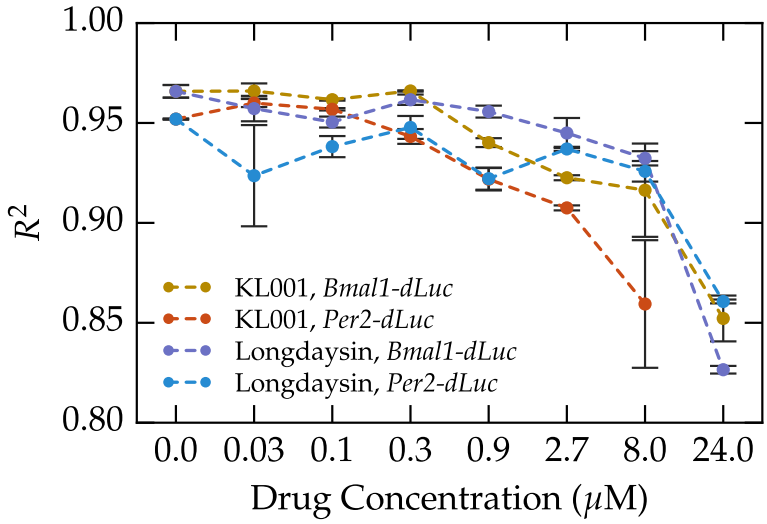


Figure 6.1: Fit quality for the dose-dependent small molecule screens. The bioluminescence rhythms in both reporter systems were well-described by a damped sinusoid. As the molecules were toxic to the reporter cells at high concentrations, fit quality declined with increasing dosage. Only fits with $R^2 > 0.8$ were kept for further analysis.

each Ω, a damped sinusoid was fit to the population-averaged state trajectory. An R^2 value was calculated for each fit, taking into account all eight state variables.

Fitting the volume parameter

We calculated an average experimental damping rate of $d = 0.0151$ from the $0\mu M$ bioluminescence trajectories for both KL001 and longdaysin. *In silico* damping rates were calculated for logarithmically spaced values of $\Omega \in (100, 500)$. Ten independent groups of $1,000$ oscillators were simulated for each Ω, from which the means and standard errors were found. Simulations in which $R^2 < 0.90$ were removed from further analysis. A weighted least-squares regression (using statsmodels, `http://statsmodels.sourceforge.net/`) was performed for $\log d$ vs. $\log \Omega$, using the $\log$ SEM of each measurement as a regression weight (figure 6.2). The best fit

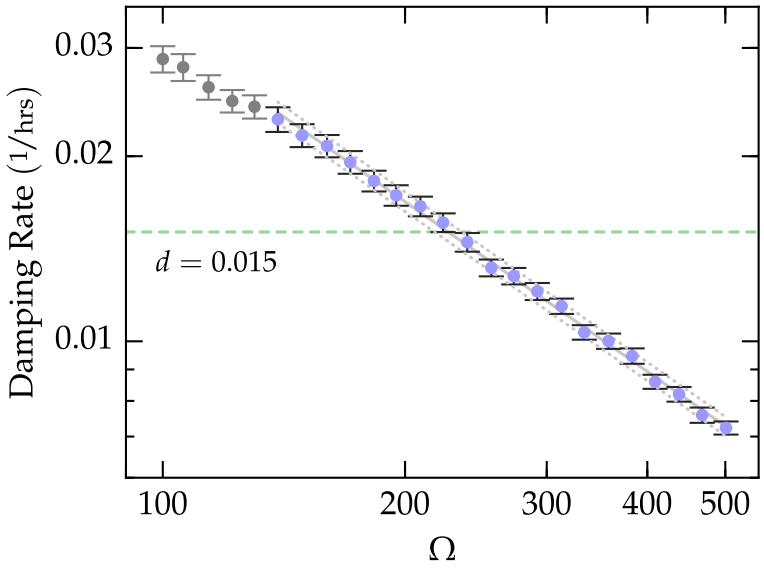


Figure 6.2: Calibration curve for fitting the volume parameter to experimental data. The model's volume parameter was linearly related to the population-averaged damping rate on a log-log scale. Error bars represent the standard error of the mean, calculated by 10 independent replicates for each volume. Points shown in gray had an average $R^2 < 0.9$ and were excluded from the linear regression. Solid and dashed grey lines indicate the mean and 95% confidence intervals of the linear regression, respectively.

was found to be $\Omega = 226.3 \pm 9.0$, and was used for subsequent model simulations.

Parameter knockdown experiments

We replicated the effects of the small molecules KL001 and longdaysin mathematically through the reductions of the $vdcn$ and $vac1p$ parameters, respectively (model 1). Knockdown simulations were performed with 20 values of each parameter, linearly spaced between 100% and 15% of their nominal value. Similar to the volume calibration simulations, 10 independent populations of $1,000$ oscillators were simulated from an initially entrained state. Means and standard errors in period and damping rate were calculated from each population-averaged trajectory. Simulations in which $R^2 < 0.90$ were removed from further analysis.

6.2.5 Fitting the genome-wide siRNA screen

We analyzed the data and annotations for the 111,743 wells ($T_s = 2,\ N = 72$) in the Zhang *et al.,* 2009 screen [22]. Fits for which the $R^2 < 0.80$ were discarded. The natural logarithm of the amplitude was used, since it more closely resembled a normal distribution and was on a similar scale to the damping rate. Plate to plate variation, as shown in figure 6.3, was more severe than variation on a well-to-well basis, figure 6.4. Parameters were therefore normalized on a plate-by-plate basis using a robust z-score [142]:

$$z_{R,i} = \frac{p_i - M(p_i)}{M\left(|p_i - M(p_i)|\right)}, \qquad i \in \{0, \ldots, \mathcal{P} - 1\}$$

where $M(\cdot)$ denotes the median of a vector, and p_i contains all the points in one plate, and $\mathcal{P}$ is the number of plates. We removed outlier points prior to calculating the moments of the distributions, Pearson correlation coefficients, and performing the multivariable linear regression. Outliers were defined as points that contained a z-metric (in either period, phase, amplitude, or damping rate) with an absolute value greater than eight. We chose the "control" wells to be those that contained no siRNA, as these proved to be more numerous than those containing reference siRNA perturbations and were clustered similarly to the highest-density regions of the perturbed fits.

Detecting outlier perturbations

We found the average response to each gene perturbation by grouping the perturbed dataset by target gene ID (using pandas, `http://pandas.pydata.org/`). A Hotelling's T^2 test was used to determine whether the means of each gene knockdown was

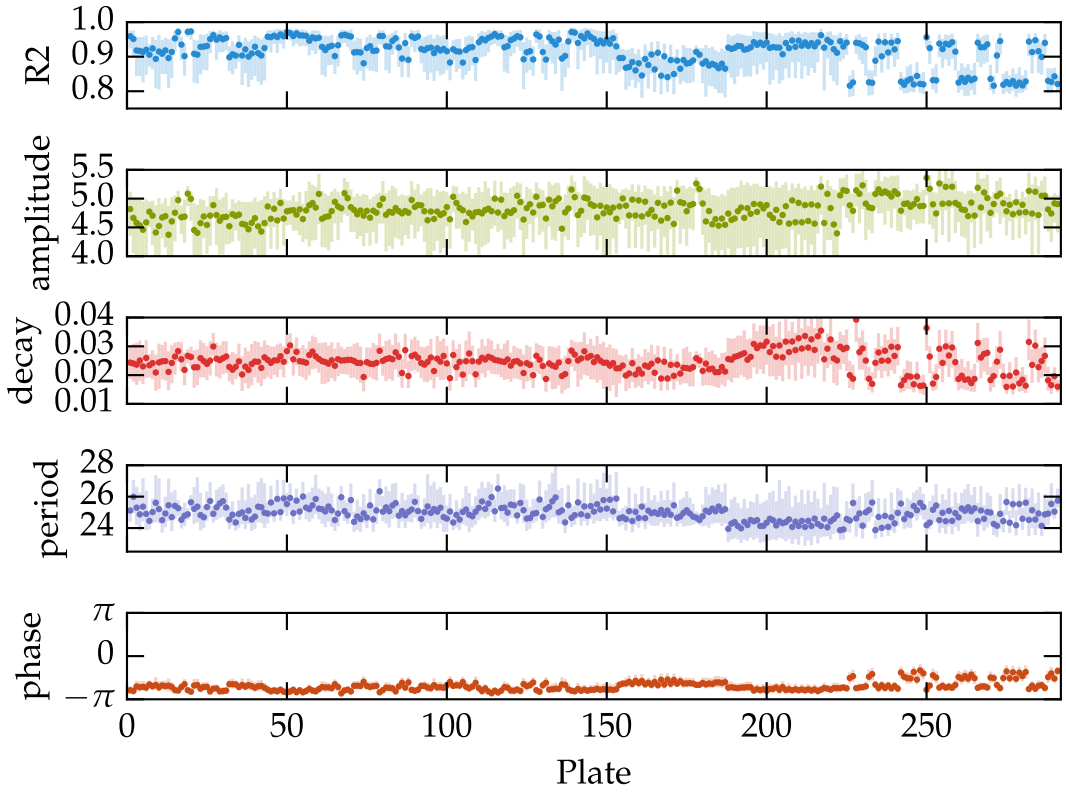


Figure 6.3: Plate-to-plate variation of fitted parameters in the Zhang *et al.*, 2009 genome-wide siRNA screen. Dots indicate the median of each plate, with lines extending from the 5^{th} to 95^{th} percentile. While parameter fits were of similar magnitude for all plates, some inconsistencies were present. In order to accurately compare perturbations and controls between plate experiments, we normalized fitted parameters on a plate-by-plate basis.

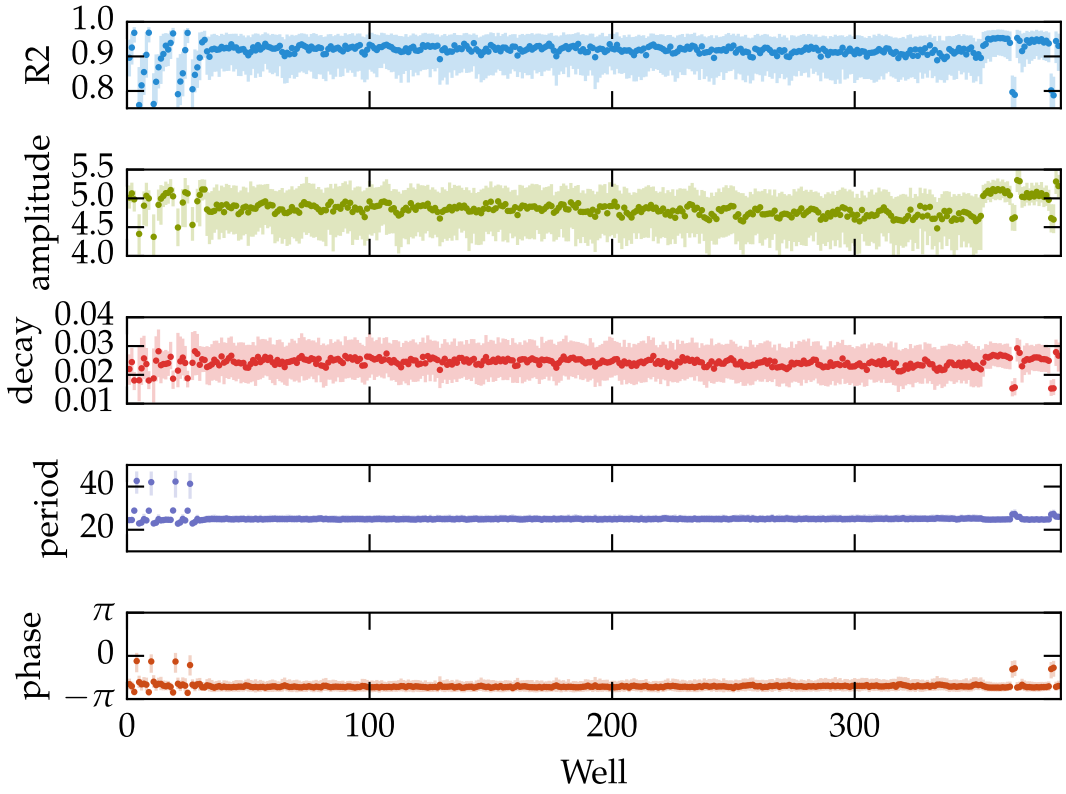


Figure 6.4: Well-to-well variation in the Zhang *et al.*, 2009 genome-wide siRNA screen. Similar to figure 6.3, dots indicate the median value of each parameter in each well, with lines extending from the 5th to 95th percentile. Well position did not seem to affect the fitted values, particularly in the middle regions that contained the siRNA knockdown library. Wells on either end showed significant variation, but these are likely due to the spotting of long- and short-period controls in the same position on each plate.

significantly different from the control population. A robust covariance estimator was used to find the location and covariance of the control distribution (using scikit-learn). Because the control distribution ($n = 11,253$) is much larger than that for any particular gene ID ($n \approx 4$), the effect of the perturbed sample on the pooled covariance was neglected.

6.2.6 Software

All computations were performed using Python. Code used to perform the analysis and produce the figures in this chapter can be found online at `https://github.com/pstjohn/decay-methods`.

6.3 Results

6.3.1 Higher noise results in faster damping in population-level rhythms

We first sought to determine whether changes to single-cell stochastic noise affect population-level damping rate. To do this, we calculated noise characteristics from experimental data on individual PER2::LUC fibroblast cells [138]. Cells were sorted into two groups, a low-noise group and a high-noise group, based on the relative high-frequency noise, period variability, and amplitude variability present in each trace. Example rhythms from cells in both groups are shown in figure 6.5A. Because the cells were not synchronized at the start of the recording, this effect is replicated *in silico* by shifting each series in time to align their start phases. Population-level bioluminescence traces were then found by averaging the cellular PER2::LUC signal in each group. Both populations displayed averaged rhythms that resembled a damped

sinusoid, similar to those seen in bioluminescence recordings of entire cell cultures. Fitting the averaged expression of each group with a damped sinusoid revealed that the low-noise group also had a lower damping rate (figure 6.5B). The significance of this difference was confirmed via a bootstrap analysis (figure 6.5C), where cells were randomly assigned in each bootstrap trial to either the low-noise or the high-noise group.

6.3.2 Clock perturbations can change single-cell stochastic noise

We next demonstrate that perturbations to the transcriptional oscillator are capable of altering single-cell level stochastic noise. The actions of small-molecule circadian modulators KL001 and longdaysin are well-characterized, and are known to affect circadian period and amplitude in a dose-dependent fashion [96]. By fitting experimental data on the population-level responses to increasing dosages of each molecule with a damped sinusoid, we show that KL001, but not longdaysin, increases damping rate in a dose-dependent fashion (figure 6.6). This change in damping rate is consistent across both reporter systems (*Bmal1-dLuc* and *Per2-dLuc* U2OS cells), indicating it is a fundamental property of the overall gene regulatory network.

To lend further support to the conclusion that the dose-dependent changes in damping rate from KL001 is due to changes in single-cell noise characteristics, we employed a model of circadian rhythms previously used to explain the effects of both small molecule perturbations [96], summarized in model 1. In order to capture changes to noise characteristics, we first converted the model to a stochastic biochemical system. Population-level rhythms were generated by averaging the trajectories of $1,000$ individual, noninteracting oscillators. The only free parameter in converting the existing deterministic model to a stochastic one is the cell volume, which was determined by fitting the observed population-level damping rate to that of the

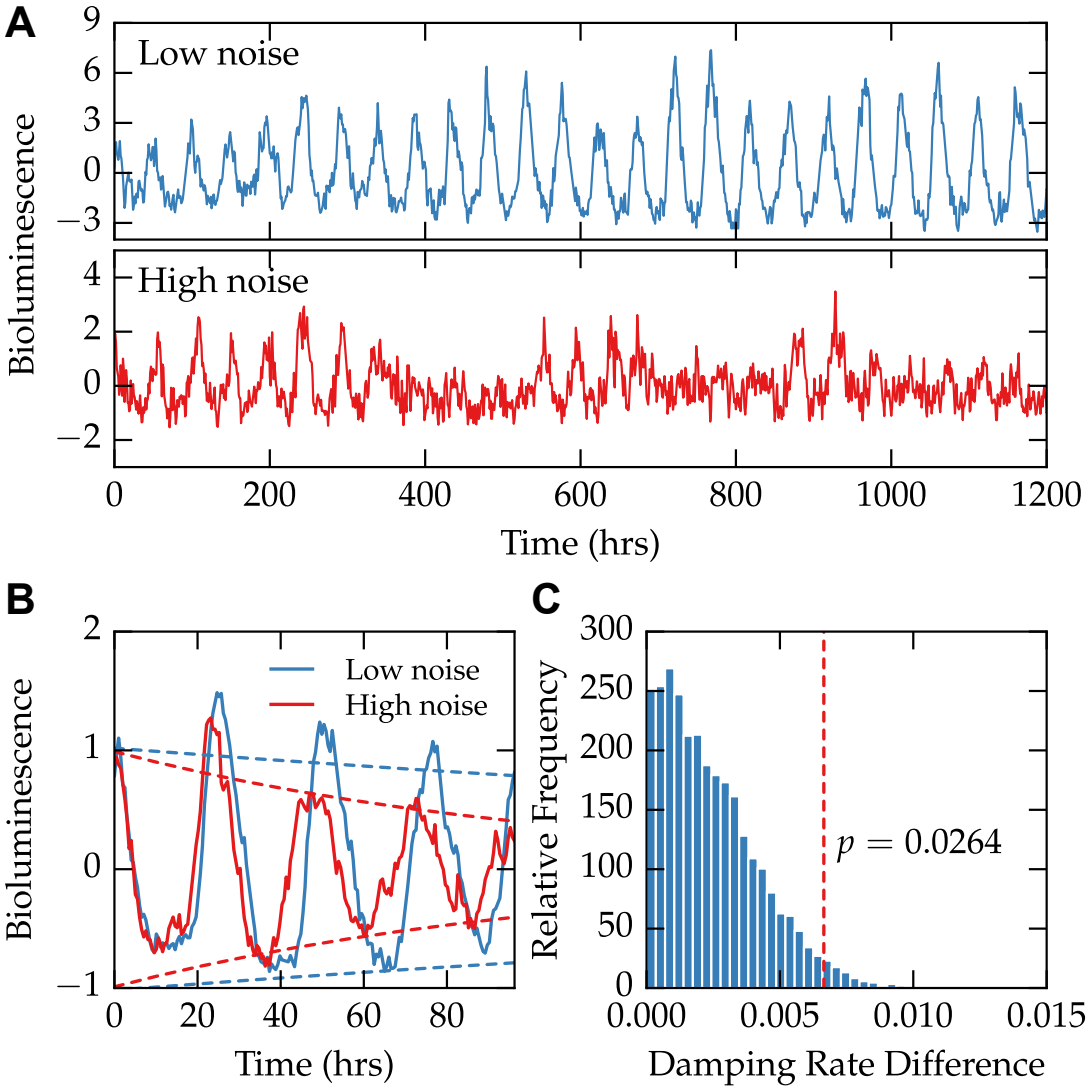


Figure 6.5: Single-cell bioluminescence recordings show that higher stochastic noise results in faster damping at the population-level. Data on the bioluminescence of single-cell fibroblasts was taken from Leise *et al.*, 2012 [138]. (**A**) Cells were sorted into two groups depending on their degree of stochastic noise. An example trace from each of the two groups is shown, demonstrating different levels of noise present in the dataset. (**B**) After artificially synchronizing each cell, we calculate averaged bioluminescence rhythms of each group (solid lines). A damped sinusoid fit to both groups reveals a difference in damping rate, demonstrated by fitted envelope functions ($\pm A \exp -dt$, dashed lines). (**C**) The observed absolute difference in damping rate was shown to be significant ($p = 0.0264$) using $10,000$ bootstrap trials.

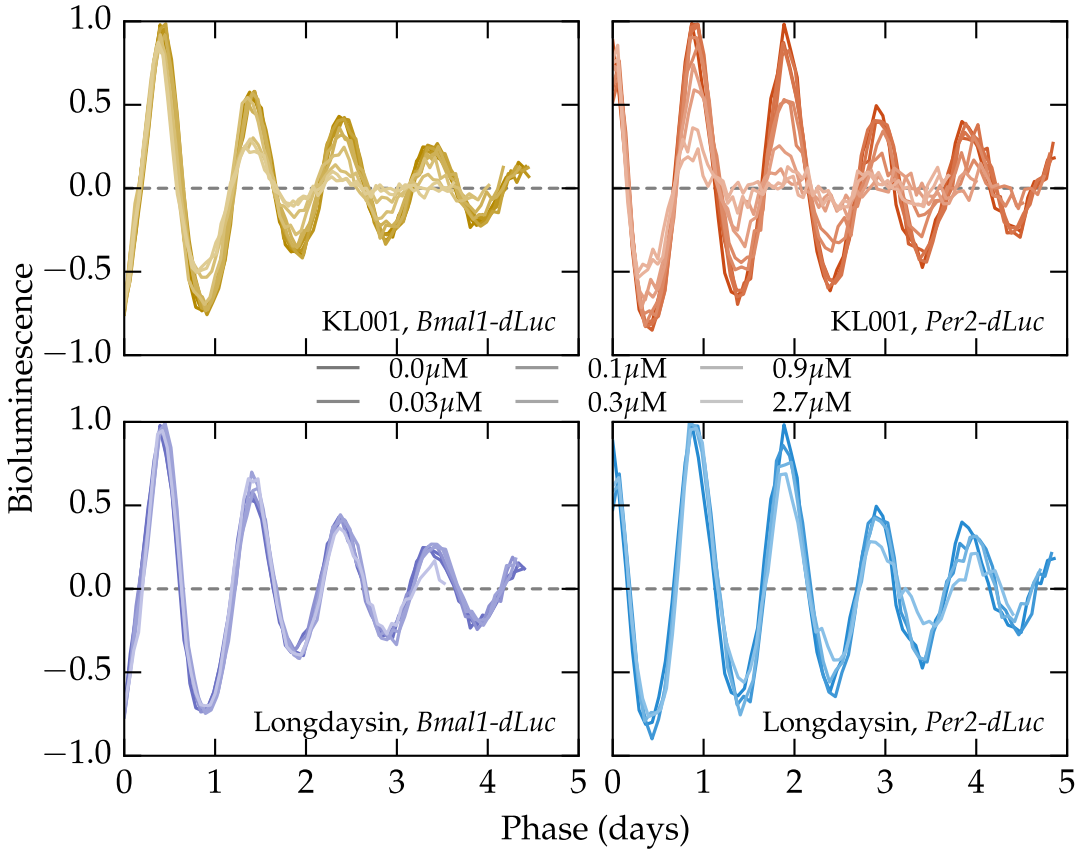


Figure 6.6: Experimental data on the dose-dependent effects of small molecules KL001 and longdaysin. Data on cultured circadian reporter cells was taken from Hirota *et al.*, 2012 [49]. Detrended bioluminescence signals from the two reporter systems and two small molecules are shown normalized by the fitted amplitude, period, and phase. The normalized bioluminescence highlights the dose-dependent change in damping rate seen with the KL001 application (top), but not with longdaysin (bottom).

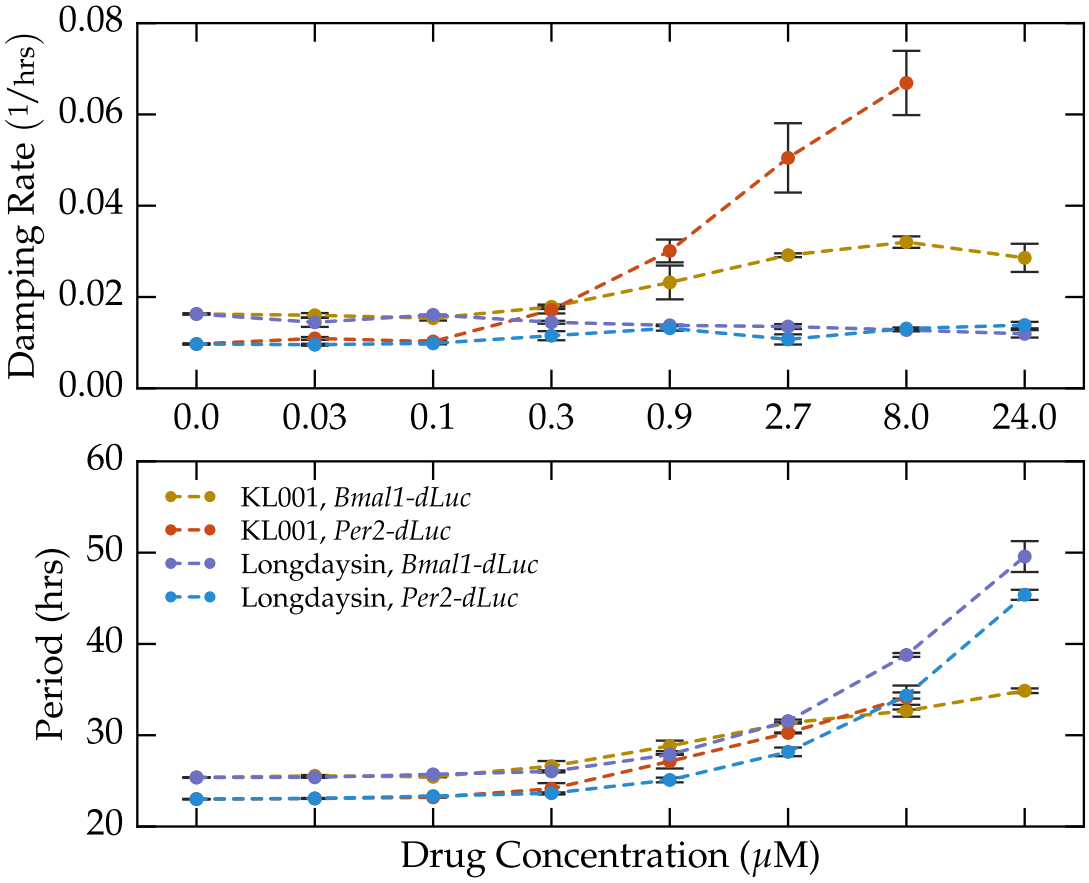


Figure 6.7: Small molecule modulator KL001 increases damping rate in a dose-dependent fashion. Quantification of the dose-dependent change in damping rate caused by small molecule modulators. While both molecules lead to a dose-dependent increase in period, only KL001 shows a reliable change in damping rate.

experimental control traces (figure 6.2).

The model was then used to predict the effects of KL001 and longdaysin administration on single-cell noise and population damping rates (figure 6.8). Reductions in parameters previously attributed to the activities of each small molecule caused dose-dependent changes in period and damping rate at the population-level that closely matched experimental results (figure 6.9). As the model includes no cell-to-cell communication, this difference is manifested solely by changing the noise characteristics of individual cells.

6.3.3 Genome-wide effects of siRNA knockdown on single-cell stochastic noise

Unlike from using single-cell imaging, inferring stochastic noise from the desynchronization rate of population-level recordings can be applied to existing and future high-throughput circadian screens. We demonstrate the insights that can be gained from such an approach by analyzing the publicly available genome-wide siRNA screen from Zhang *et al.*, 2009 [22]. The results of fitting a damped sinusoid to each of the 111,743 bioluminescence trajectories is shown in figure 6.10 and figure 6.11, in which 86% of fits had an $R^2 > 0.8$. Since sinusoidal parameters can only be confidently inferred for fits in which the R^2 was sufficiently high, wells were removed from further analysis if $R^2 < 0.8$. Additionally, of the fits with a high R^2 value, only a small minority (0.1%) had a negative damping rate. This trend lends further support to the assumption that intercellular synchronization is unlikely in cultured U2OS cells.

We next checked how parameters varied on a plate-to-plate and well-to-well basis. Well-to-well variation was relatively absent, aside from expected variation due to

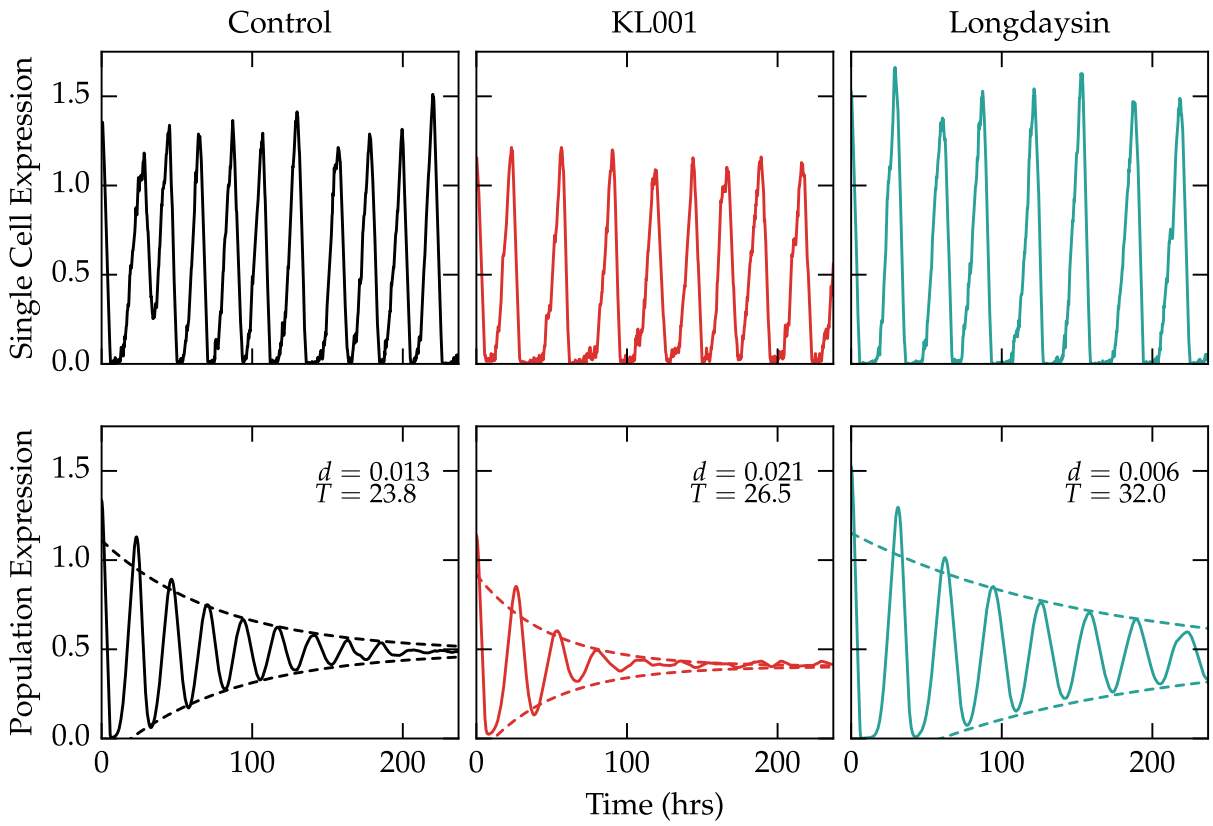


Figure 6.8: Example model trajectories. Example single-cell trajectories (top) and population-averaged trajectories (bottom, mean of 1,000 cells) of cells under various treatments. Cells with the nominal parameter set (left, black) closely match the experimental damping rate for unperturbed cells. Cells with simulated KL001 action (red, center) are noisier at the single-cell level, and show faster damping at the population-level. Cells with simulated longdaysin action show slightly more accurate single-cell rhythms, with a corresponding reduction in the population-level damping rate.

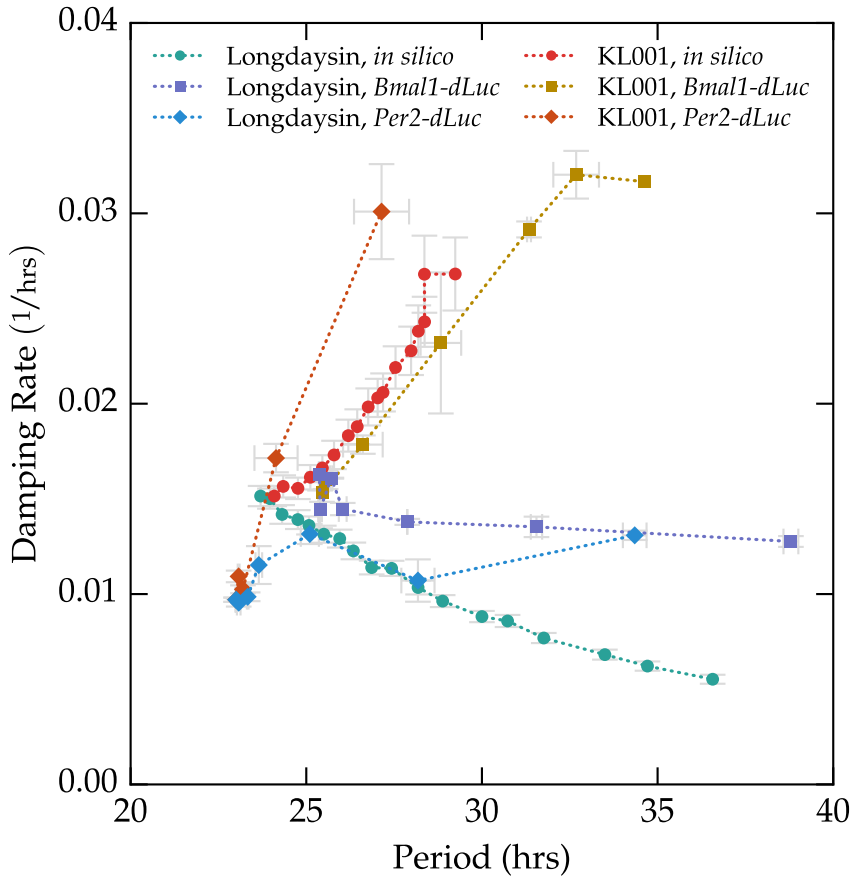


Figure 6.9: Mathematical model accurately predicts dose-dependent changes in damping rate. The model accurately predicts the general trend of period vs. damping rate for both KL001 and longdaysin perturbations. Experimental data points represent the mean of two replications at each concentration. Computational data points represent the mean of ten independent population simulations.

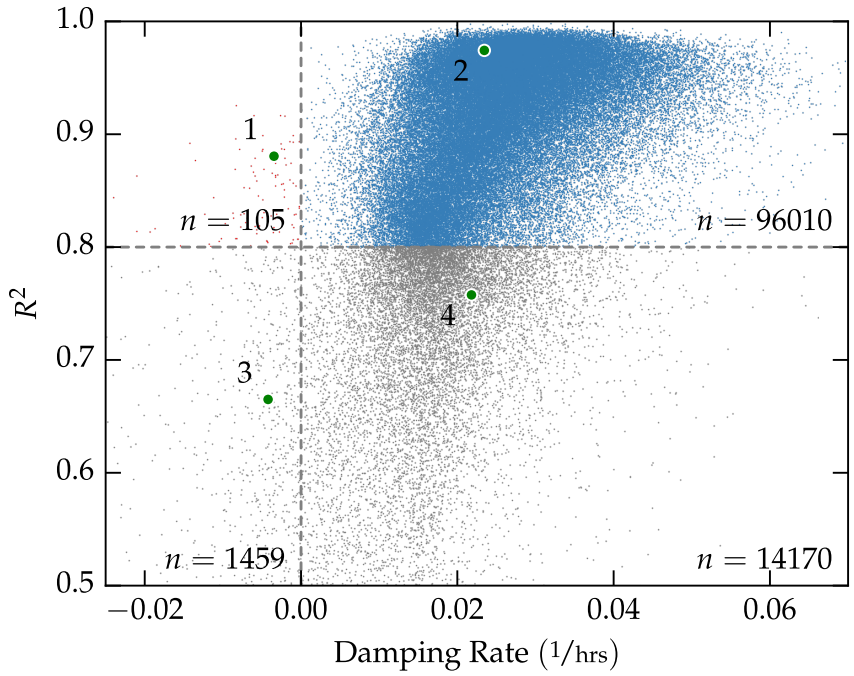


Figure 6.10: Fit quality vs. damping rate for the genome-wide siRNA screen. A plot of the 111,743 individual fits shows that the majority of fits have a high R^2 value and positive damping rate. Only fits with $R^2 > 0.8$ were kept for further analysis.

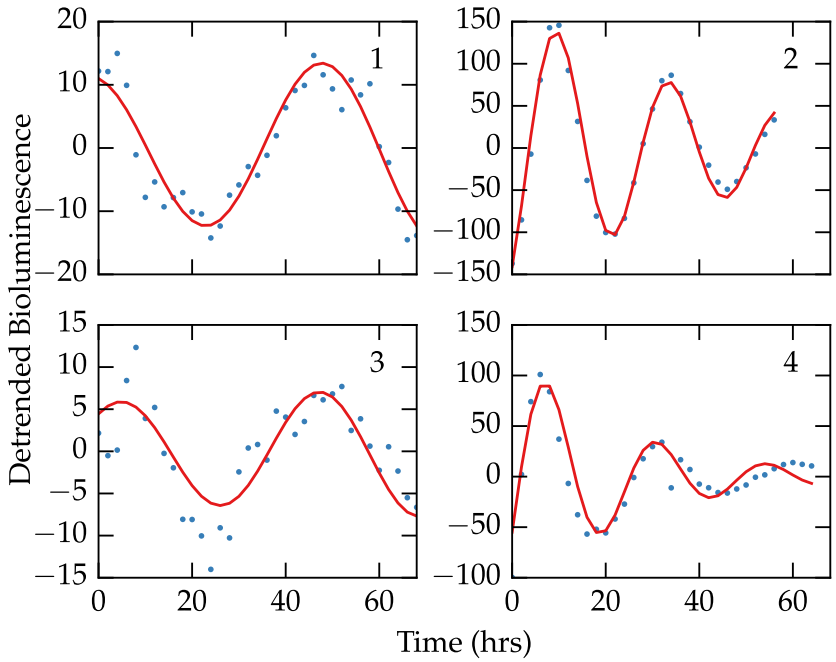


Figure 6.11: Example fitted bioluminescence trajectories. Examples chosen randomly from each of the four quadrant regions in figure 6.10. Sinusoidal parameters for fits 1-2 can be more confidently inferred than those for fits 3-4.

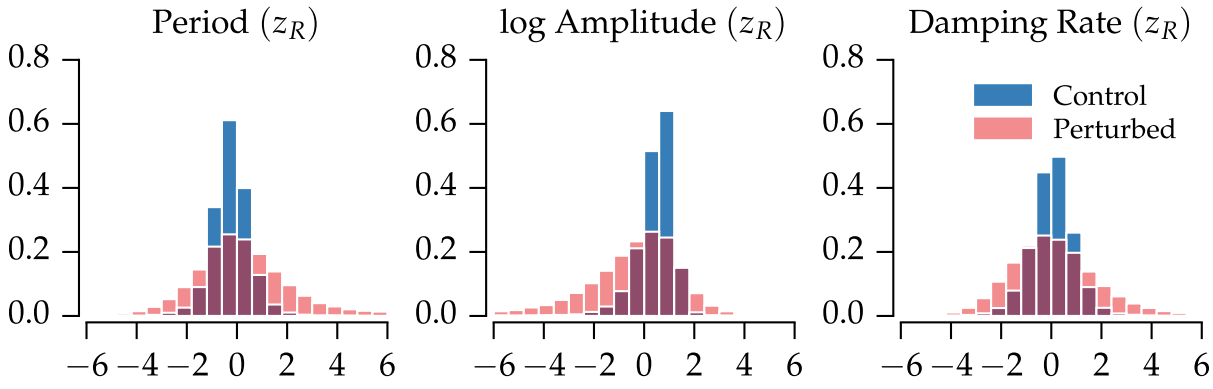


Figure 6.12: Distributions in fitted parameters for the genome-wide siRNA screen. Distributions in robust z-scores closely resemble normal distributions. For all parameters, the region of highest density is consistent between the control and perturbed populations, indicating many perturbations do not appreciably change clock dynamics.

long- and short-period controls (figure 6.4). Fits were normalized to remove plate-to-plate variation (figure 6.3) using a robust z-score [142]. Additionally, we separated wells into a "perturbed" category and "control" category, depending on whether or not the well contained an siRNA perturbation. As we show in figure 6.12, all fitted parameters displayed normal-like distributions, in which the control distributions showed much tighter clustering around the most likely values. Quantifications of the mean, variance, skewness, and kurtosis for each distribution are shown in table 6.1.

Damping rate is independent of other sinusoidal parameters

Since an siRNA's effect on period length does not effectively predict its effect on amplitude or phase (and vice-versa), we hypothesized that stochastic noise would similarly be independently affected by siRNA perturbations. Low Pearson correlation coefficients between normalized parameter distributions defend this hypothesis, with the highest correlations among variables seen between amplitude and damping rate ($\rho = 0.285$, table 6.2). Additionally, a multivariate linear regression of damp-

	T		$\ln A$		d	
	C	P	C	P	C	P
Mean	−0.234	0.187	0.443	−0.343	0.043	0.090
Variance	0.599	3.313	0.605	3.074	0.771	2.850
Skewness	0.153	0.367	−1.823	−0.580	−0.107	0.371
Kurtosis	3.772	0.591	8.329	0.476	2.423	0.373

Table 6.1: Moments of the fitted parameter distributions (figure 6.12) after normalization and outlier removal. Parameters were normalized by subtracting the median and dividing by the median absolute deviation on a plate-by-plate basis. Control distributions had less variance, were less skewed, and were more peaked then their perturbed counterparts.

	d	$\ln A$	T	θ
d	1.000	0.285	−0.142	−0.269
$\ln A$	0.285	1.000	−0.022	−0.112
T	−0.142	−0.022	1.000	−0.113
θ	−0.269	−0.112	−0.113	1.000

Table 6.2: Correlation among normalized parameters of the high-throughput siRNA screen. Pearson correlation coefficients are relatively low between fitted parameters, indicating that changes to damping rate (and thereby stochastic noise) are not explained by changes to period, amplitude, or phase.

ing rate as a function of period, amplitude, phase, and perturbation type (control or perturbed, a categorical variable) produced an R^2 value of only 0.169 (table 6.3). These results reinforce the notion that the population-level damping rate describes an independent feature of the underlying oscillator.

The unperturbed clock lies at the Pareto frontier of high amplitude and low noise

The ability of a population of oscillators to maintain robust oscillations is a function both of its initial amplitude as well as its damping rate. A robust oscillator is therefore one with high amplitude and low stochastic noise. Scatter plots of the control

Dep. Variable:	Damping Rate	**R-squared:**	0.169
Model:	OLS	**Adj. R-squared:**	0.169
Method:	Least Squares	**F-statistic:**	4782.
Date:	Wed, 11 Feb 2015	**Prob (F-statistic):**	0.00
Time:	16:26:22	**Log-Likelihood:**	$-1.7248e+05$
No. Observations:	94053	**AIC:**	$3.450e+05$
Df Residuals:	94048	**BIC:**	$3.450e+05$

	coef	**std err**	**t**	**P>|t|**	**[95.0% Conf. Int.]**
intercept	-0.0370	0.014	-2.572	0.010	$-0.065, -0.009$
amplitude	0.2375	0.003	86.282	0.000	$0.232, 0.243$
period	-0.1521	0.003	56.798	0.000	$-0.157, -0.147$
phase	-0.2354	0.003	85.598	0.000	$-0.241, -0.230$
perturbation type	0.3197	0.015	20.664	0.000	$0.289, 0.350$

Omnibus:	9769.391	**Durbin-Watson:**	1.876
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	18459.719
Skewness:	0.697	**Prob(JB):**	0.00
Kurtosis:	4.664	**Cond. No.**	8.34

Table 6.3: Multivariable linear regression results. Fit statistics that demonstrate the effect of each other fitted parameter on damping rate. Of particular note is the perturbation type categorical variable, which demonstrates that the presence of siRNA perturbation increases damping rate on average, controlling for changes in other variables. Higher amplitude is also correlated with higher damping rate. However, in total damping rate is poorly predicted by the other fitted variables ($R^2 = 0.169$), indicating it describes an independent oscillatory feature.

and perturbed distributions are shown in figure 6.13A-B, which indicate that few outlier points reside in the lower damping rate, higher amplitude quadrant. In order to find perturbations that confidently change robustness, we grouped the siRNA perturbations by target gene and performed a two-sample Hotelling's T^2 test against the control population. The resulting significant gene perturbations (75% of all genes) are shown in figure 6.13C. Quantifying the distribution in outliers by quadrant, it is clear that only a small minority of perturbations (3.3%) simultaneously increase amplitude and decrease stochastic noise. It therefore appears that the unperturbed clock optimizes some trade-off between high amplitude and low stochastic noise, such that it is unlikely the knockdown of a single gene will make oscillations more robust.

6.3.4 Effect of cell heterogeneity

Experimental work has shown that cell-autonomous fibroblast cells have a distribution of free-running periods [138]. Indeed, prior to the availability of single-cell data, studies explored the possibility that differences in mean periods served as the mechanism behind population-level damping [143]. While it is true that the dephasing of a group of oscillators can be caused by both variance in the mean period as well as cycle-to-cycle variability, in the preceding analyses we neglect the effect of period heterogeneity. To motivate this assumption, we show that there is greater variance in period on a cycle-to-cycle basis than on a cell-to-cell basis (figure 6.14). Additionally, it is important to note that the dephasing kinetics of these two alternative processes are distinguishable from population-level data. In dephasing due solely to stochastic noise, the probability density of finding a cell with any given phase follows a standard convection-diffusion regime, resulting in exponentially damped rhythms [122]. In dephasing due solely to period heterogeneity, the variance of the phase probability density broadens in time proportionally to t^2, rather than t in standard diffusion. This

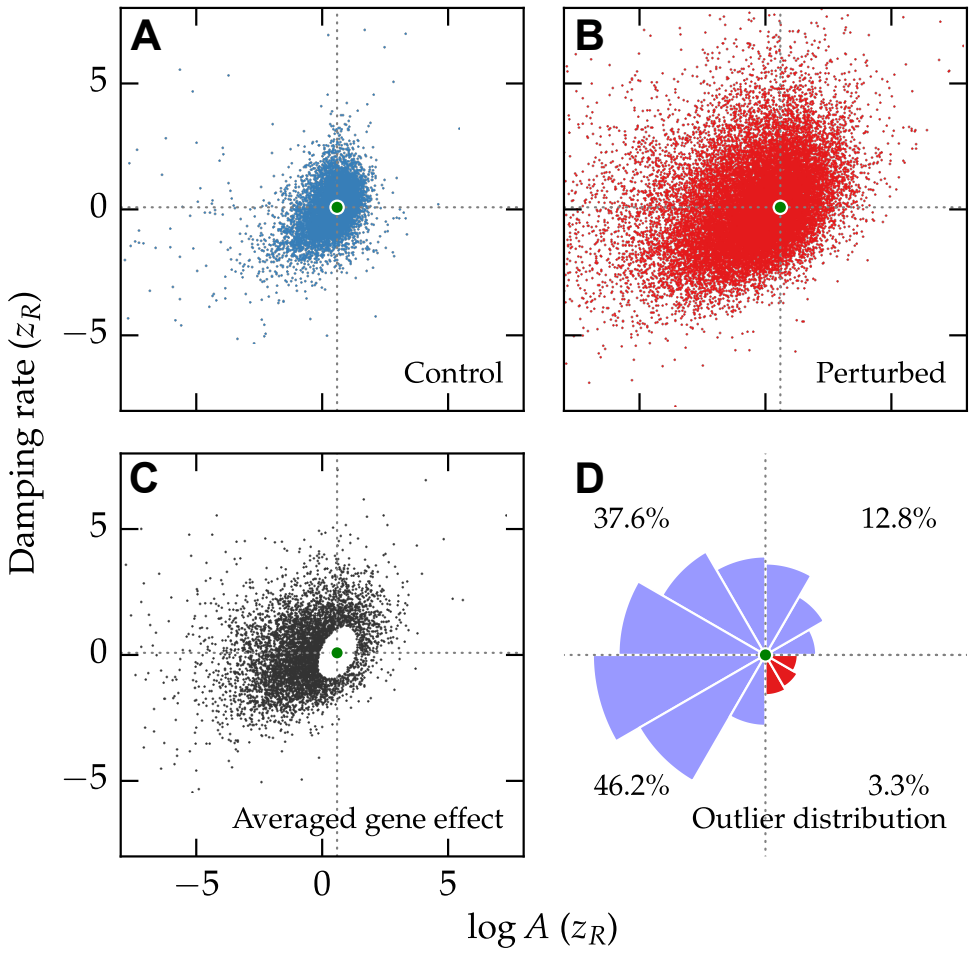


Figure 6.13: Effects of siRNA knockdowns on amplitude and damping rate. Clock robustness is a function of both amplitude and damping rate. Distributions in amplitude and damping rate for control wells (**A**) or perturbed wells (**B**) indicate that perturbations tend to shift the clock towards regions of higher damping rate or lower amplitude. Green dots in each figure indicates the mean of the control population. (**C**) Averaged effect of siRNA knockdown after grouping the perturbed population by Gene ID. Only those genes that were significantly different from the control distribution are shown (Hotelling's T^2 test, $\alpha = 0.01$). (**D**) Radial histogram of the significant gene perturbations shown in (C). The area of each slice is proportional to the frequency of perturbations away from the mean in that direction. Very few gene knockdowns result in both higher amplitude and lower damping rate (red slices, lower right quadrant).

type of anomalous diffusion results in a different damping profile, as highlighted in figure 6.15. Since the data used in this chapter is well-described by an exponentially damped sinusoid, and since a mathematical model assuming homogeneous mean periods accurately predicts the experimental damping rate changes, this assumption is likely justified. Additionally, since this chapter is mainly concerned with changes in the damping rate (rather than the absolute damping rate), it is unlikely that period heterogeneity plays a major role.

6.4 Discussion

In this chapter, we have described a method by which changes to stochastic noise can be estimated from population-level circadian bioluminescence recordings. As high amplitude circadian oscillations are important for maintaining metabolic health, many studies have sought to find small molecule candidates that increase oscillatory amplitude [48]. In the search for circadian clock therapeutics, high-throughput methods are frequently used to screen for such drugs, often neglecting potential effects on cell-autonomous noise. While we have demonstrated a method by which the effects of small molecules on noise can be inferred from high-throughput methods, we have also shown that the potential improvement of clock robustness may be limited. Amplitudes of circadian rhythms may therefore be best increased by small molecule therapies that act transiently to synchronize peripheral oscillators. While such a method would require accurate alignment of drug administration to the correct circadian phase, a recent *in silico* study has demonstrated the potential effectiveness of such an approach in improving amplitudes in peripheral tissues [122]. Finally, the ability to extract an additional biologically relevant parameter from existing datasets will likely prove useful for many studies, as it allows further differentiation between per-

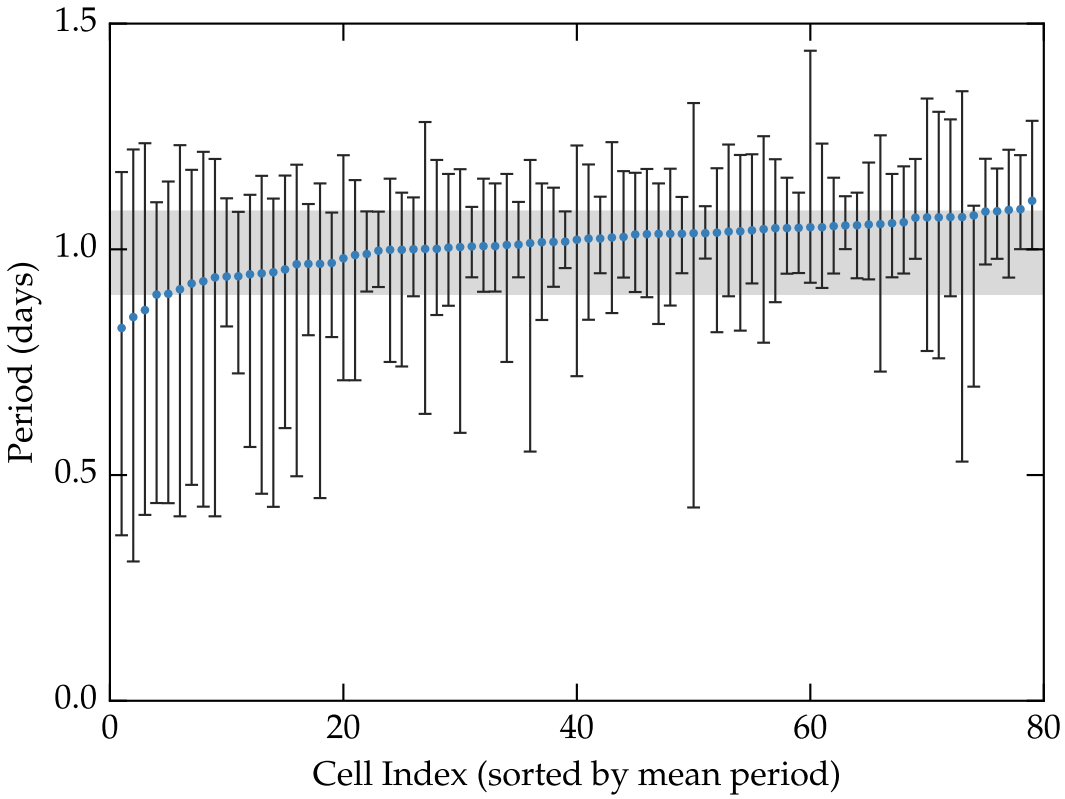


Figure 6.14: Variability in period length from cycle to cycle is greater than from cell to cell. Using single-cell fibroblast data from [138], a distribution of period lengths were calculated for each cell. Average period length for each cell is shown by a blue dot, with cells sorted from lowest to highest average period. Error bars extend from the 5^{th} to 95^{th} percentile period within each cell. The gray shaded region extends from the 5^{th} to 95^{th} percentile of the average period lengths for the entire population. 87% of cells show greater variability from cycle-to-cycle than the overall period heterogeneity.

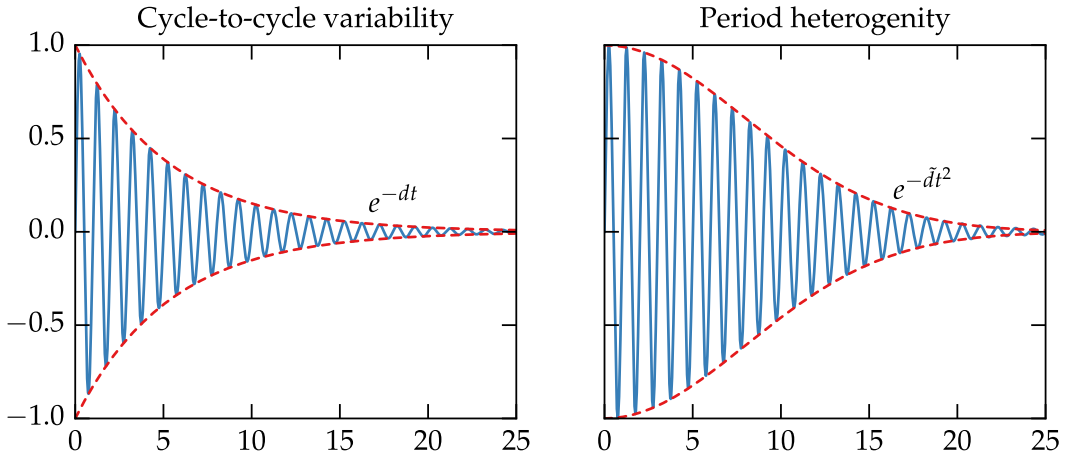


Figure 6.15: Period heterogeneity and phase diffusion display different damping profiles. (*Left*) A simulated population of homogeneous, noisy oscillators dephases due to cycle-to-cycle variability. The envelope of the averaged expression is proportional to $t9$ (*Right*) A deterministic, heterogeneous population dephases due to different free-running periods. In this case, the population displays ballistic phase diffusion, in which envelope of the population-averaged expression changes proportionally to t^2.

turbations that might otherwise have identical effects on bioluminescence rhythms.

Chapter 7

Conclusions and future work

The preceding chapters discussed several research threads that have been brought more or less to completion during my PhD thesis. It was relatively fortunate that the model created early in my PhD thesis, described in chapter 2, has remained current and predictive for the past few years. This model was used in some capacity in each of my publications, whether as simply an example oscillator (in chapter 3 or chapter 5), or for its biological insight (in chapter 4 or chapter 6). Another theme throughout this thesis has been the development and subsequent implementation of various computational methods or approaches. Specifically, the identifiability scheme discussed in chapter 3 was used to make the predictions described in chapter 4, and many of the ideas described in chapter 5 were used as the motivations behind the more data-driven analysis in chapter 6.

Thanks in part to my Mitsubishi Chemical fellowship, each chapter in this thesis largely follows from the work in the preceding it. Throughout this process, many interesting possibilities for further analysis were considered. This chapter describes several such opportunities for future studies.

7.1 Experimental validation of the amplitude response to metabolic stimulation

One of the key motivations for the development of the amplitude response curve metrics was applying the methods to metabolic perturbations. Several studies have highlighted the importance of rhythmic food intake in maintaining high amplitude rhythms in peripheral tissues [15, 144]. Additionally, the importance of high-amplitude peripheral oscillations in maintaining metabolic health has been demonstrated multiple times [12, 15]. It is therefore likely that signals coming from rhythmic changes in metabolic state entrain peripheral oscillators and give them a temporary boost in amplitude.

In particular, we were motivated by results from Yamajuku *et al.*,, 2012, which shows that a temporary dose of insulin not only changes the phase of cultured hepatocytes, but also appears to have an effect on the amplitude of the oscillations as well (figure 7.1). This effect, while not described in the original paper, seems to be phase-dependent: insulin delivered at late times (T12-20) seems to have a small to negative effect on the amplitude, while that delivered at early times (T0-T8) has a positive effect on the amplitude. This type of behavior would be expected if insulin indeed conveyed a "fed" state to the circadian oscillator, as amplitudes would only be boosted when meals arrived at the proper phase.

Ultimately, however, the data presented in [145] did not prove to be sufficient for a complete analysis. In particular, we noted that the amplitude response due to the application of the vehicle (comparing black lines before and after the pulse, figure 7.1) seems to be as significant as the amplitude change due to the application of insulin. This is not surprising, as peripheral oscillators as acutely sensitive to

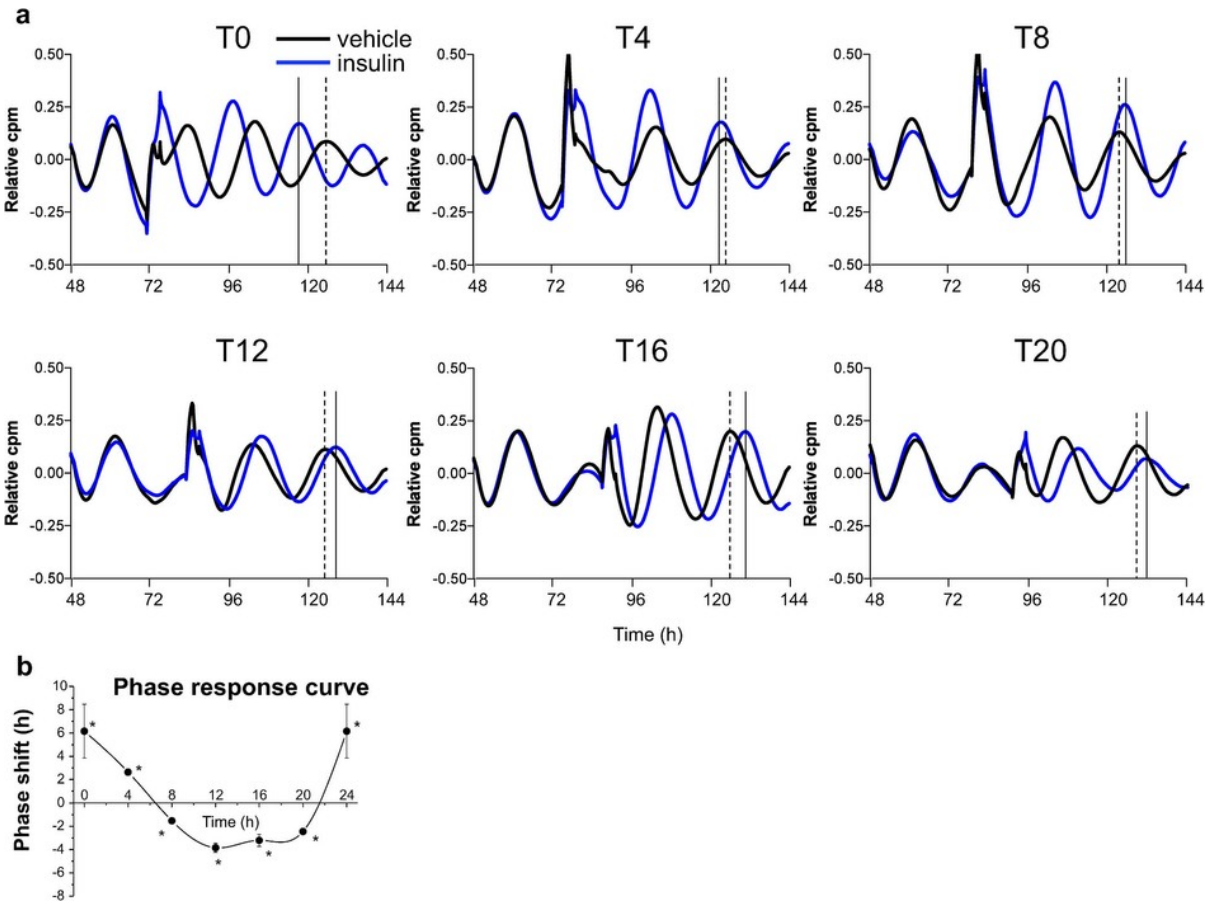


Figure 7.1: Effect on temporary insulin dosages on hepatocyte circadian gene expression. The introduction and subsequent washing out of insulin in cultured hepatocytes leads to a phase-response curve (B), but perhaps more interestingly also seems to have an effect on the amplitude of circadian oscillations. Figure reprinted from [145].

changes in temperature and culture medium, which is often enough to resynchronize the population of oscillators [26].

The recreation of these results, using careful methods to prevent perturbations from the experimental protocol may lead further insight into the interplay between circadian and energetic regulatory networks. By knocking out different components of the signal transduction pathway between insulin and the core clock, it may be possible to elucidate a mechanism by which the amplitude response curves were occurring.

7.1.1 Metabolic-circadian connections

The literature on signalling networks connecting metabolic and circadian regulators is very dense. Additionally, the signalling network changes over time, since many key metabolic regulators are expressed in a circadian manner. For example energy sensors Akt and AMPK are up- or down-regulated depending on the energy and redox state of the cell [146]. Akt, which is also a direct target of insulin [147], is known to regulate mTOR [148], which is in turn known to regulate the light-mediated transcription of *Per* in the SCN [149]. Some of the key connections between insulin and core clock species are highlighted in figure 7.2

Many posttranslational regulators are also means by which cellular energy state affects circadian rhythms. AMPK is known to affect the posttranslational regulator FBXL3, a regulator of CRY stability identified as the main target of KL001 [150]. Additionally, AMPK has been shown to affect the activity of CKIδ/ϵ [91], modulators of PER phosphorylation identified as main targets of longdaysin.

Since many of the targets of insulin could be independently stimulated via small molecule activators, a follow up experiment might be to directly test the time-dependent effect of small molecule modulators on circadian amplitudes. A summary

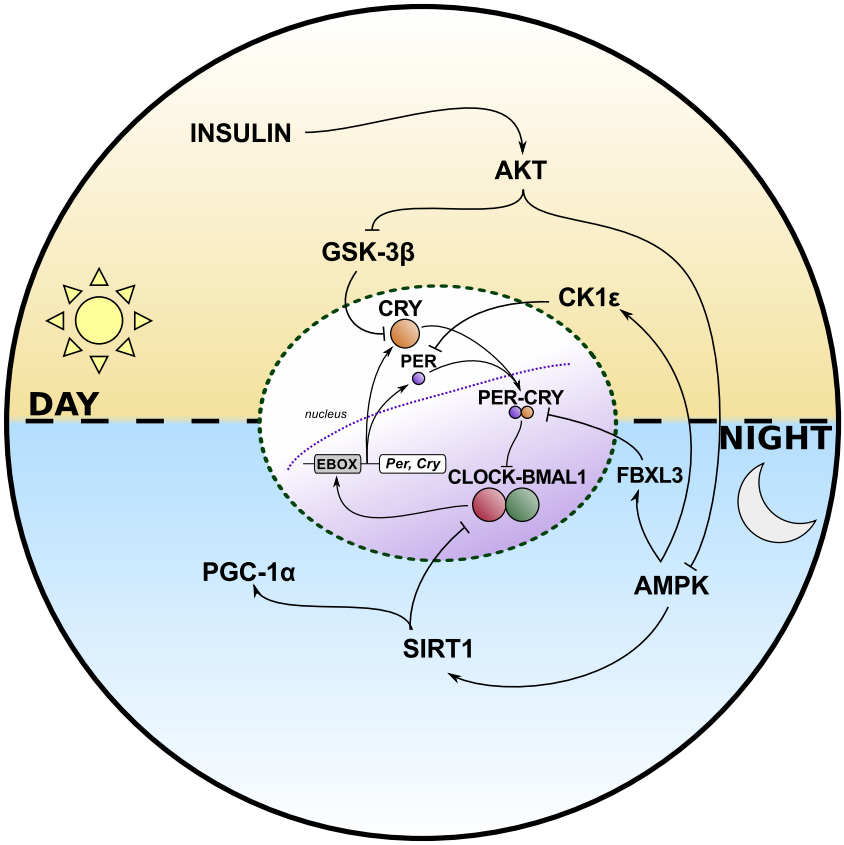


Figure 7.2: Known connections between insulin and circadian rhythms. A key step in determining potential mechanisms for the amplitude effect of insulin stimulation on circadian rhythms will be to map major signalling pathways that may be responsible for conveying information on metabolic state to the clock. Some of the connections are shown here, many of which could be independently stimulated via small molecule action.

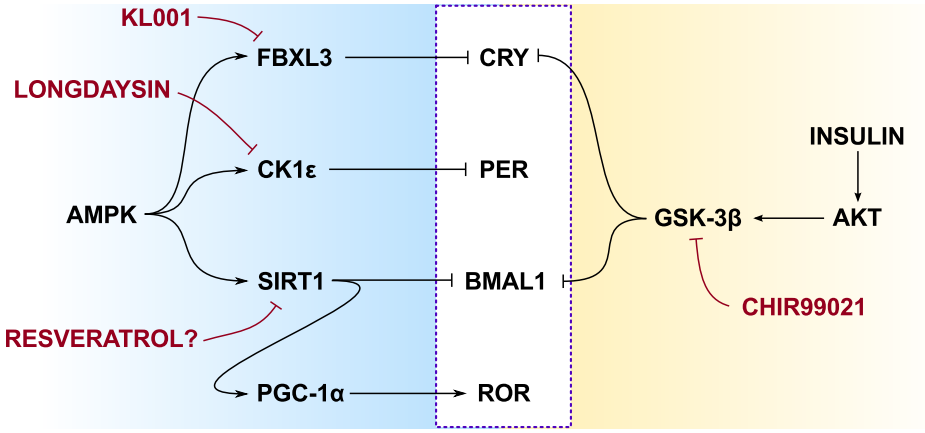
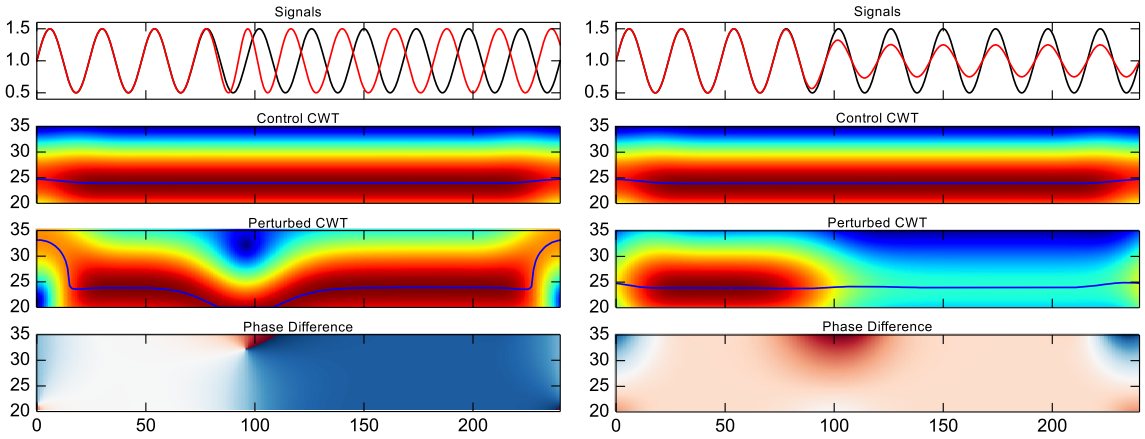


Figure 7.3: Potential small molecule candidates. Several small molecules are known to modulate intermediate species in the insulin to core clock pathway. The phase-dependent effect of these small molecules could be independently tested in order to validate experiments using insulin.

of possible small molecule modulators and pathways by which they might affect the core clock is shown in figure 7.3

7.1.2 Continuous wavelet methods

Several methods could be used to determine changes in amplitude or phase from experimental data. One of method for understanding nonstationary sinusoidal signals is the continuous wavelet transform (CWT) [151], which maps the correlation between a wavelet function and the data at a variety of scales (periods) and times. The resulting CWT is useful for visualizing changes in phase or amplitude, as shown in figure 7.4. Such a tool could be used to compare the phase-dependent changes in phase and amplitude following insulin or small molecule application under a variety of conditions.



(a) Phase Change **(b)** Amplitude Change

Figure 7.4: Application of the continuous wavelet data to example data. The continuous wavelet transform can be used to visualize variations in period, amplitude, or phase of a sinusoidal signal. In **(a)**, a reference signal (black) is compared to a perturbed signal (red) with a change in phase (top). In the middle two plots, heat maps indicate the amplitude of the CWT, which remains comparable across the two signals. In the bottom plot, the difference in phase between the two signals is highlighted by a strong transition from white to blue. In **(b)**, a change in amplitude is highlighted by a reduction in the corresponding amplitude of the CWT, with little change in phase.

7.1.3 Summary

While potentially an exciting area of research, the amplitude response characterization work poses significant experimental difficulty. In particular, since medium changes have such a profound effect on cultured cells, it is important to perform as few disturbances as possible during the incubation period. Unlike with studies using light as a temporary perturbation, it is likely impossible to remove exogenous inputs after their application, a requirement for finding their phase-dependent effect. Perhaps with a more cleverly designed experimental protocol, such as coupling the induction of insulin to a light-activated promoter, such an experiment would be possible to perform accurately. Until such time, however, the amplitude response theory discussed in chapter 5 will likely be primarily applicable to experimental work involving light-mediated changes.

7.2 Further extensions to the population phase response theory

In chapter 5, we pull together definitions from a variety of different sub-fields to create a model of a population of noisy oscillators, each following their own limit cycle dynamics. In several cases, there were further analytical treatments that could have solved in greater detail. Some of these analytical derivations would likely be necessary for future studies considering noisy systems, in which the PRC can not be assumed to be a continuous function. In this section, we present some unfinished derivations on phase populations undergoing temporary perturbations. We hope these derivations will prove useful for future studies looking to model the responses of experimental systems to light or pharmacological modulation.

7.2.1 Infinitesimal population PRCs and ARCs for limiting cases

In chapter 5, amplitude response curves at the population-level are calculated numerically by performing a change of variables of the initial probability density function $p(\theta)$ to the probability after perturbation, $\hat{p}(\theta)$. Infinitesimal response curves at the population level can be calculated numerically. Phase and amplitude response curves for the population level were previously defined as:

$$z := \int_0^{2\pi} e^{i\theta} p(\theta, \tilde{t})\, d\theta \tag{5.28}$$

$$\hat{z} := \int_0^{2\pi} e^{i\theta} \hat{p}(\theta, \tilde{t})\, d\theta = \int_0^{2\pi} e^{ig(\theta)} p(\theta, \tilde{t})\, d\theta, \tag{5.29}$$

$$\Delta\bar{\theta} = \angle\hat{z} - \angle z \tag{5.30}$$

$$\Delta\rho = |\hat{z}| - |z|. \tag{5.31}$$

It would be preferable if infinitesimal curves,

$$\lim_{\Delta x \to 0} \frac{\Delta\bar{\theta}}{\Delta x}$$

be calculated analytically in the differential limit, and might lead to more efficient insight into model kinetics and the relationship between the PRC and population synchrony.

As $\Delta x \to 0$, the phase transition function $g(\theta)$ can be approximated as

$$g(\theta) \to \theta + \frac{d\theta}{dx}\Delta x.$$

While its unlikely that the infinitesimal PRC and ARC could be calculated for an arbitrary $p(\theta)$, it may be possible to calculate them for the limiting cases in which

the oscillators were perfectly synchronized, $p(\theta) \to \delta(\theta - \mu)$, or perfectly disperse, $p(\theta) \to \mathcal{U}$.

Starting first with the case where $p(\theta) \to \delta(\theta - \mu)$, we can find $\hat{z}$ directly:

$$\hat{z} = \int_0^{2\pi} e^{ig(\theta)} \delta(\theta - \mu)\, d\theta = e^{ig(\mu)} = e^{i(\mu + \frac{d\theta}{dx}\big|_\mu \Delta x)},$$

In which $\frac{d\theta}{dx}\Big|_\mu$ represents the infinitesimal phase response curve evaluated at phase μ. After finding $z = e^{i\mu}$ similarly, the phase response curve can be found from:

$$\begin{aligned}
\lim_{\Delta x \to 0} \frac{\Delta \bar{\theta}}{\Delta x} &= \lim_{\Delta x \to 0} \frac{\arg e^{i(\mu + \frac{d\theta}{dx}\big|_\mu \Delta x)} - \arg e^{i\mu}}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\mu + \frac{d\theta}{dx}\Big|_\mu \Delta x - \mu}{\Delta x} \\
&= \frac{d\theta}{dx}\Big|_\mu
\end{aligned}$$

Which is expected, as the entire population experiences the same phase change. The amplitude response is similarly easy to find, if uninformative:

$$\begin{aligned}
\lim_{\Delta x \to 0} \frac{\Delta \rho}{\Delta x} &= \lim_{\Delta x \to 0} \frac{\left| e^{i(\mu + \frac{d\theta}{dx}\big|_\mu \Delta x)} \right| - \left| e^{i\mu} \right|}{\Delta x} \\
&= 0
\end{aligned}$$

This result follows from the infinitely synchronized population, as are not able to respond differently to the incoming pulse. A more informative solution would use a power series expansion as the variance of the initial population approached zero, finding the leading order terms of the amplitude response curve.

For the opposite limit, $p(\theta) \to \mathcal{U}$, intuition tells us that the phase response curve would be type 0, as a perturbation at any phase would induce the same new mean

phase in the population (and thus lead to an infinite iPRC). The amplitude response curve would similarly be a single value for all phases, representing the magnitude of the new induced phase. Working through the analytical derivations, however, is more complicated. For the unperturbed population,

$$z = \frac{1}{2\pi}\int_0^{2\pi} e^{i\theta} d\theta = 0.$$

For the perturbed population,

$$\begin{aligned}\hat{z} &= \frac{1}{2\pi}\int_0^{2\pi} e^{ig(\theta)} d\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi} e^{i(\theta + \frac{d\theta}{dx}\Delta x)} d\theta\end{aligned}$$

As $\frac{d\theta}{dx}$ is a function of θ, the above integral would likely require a series solution for small Δx. Finding the absolute values and arguments of the series solution would then pose additional hurdles.

7.2.2 Phase Response distributions

Stochastic noise at the single-cell level plays an important role in how cells respond to temporary perturbations. In the analytical derivations of chapter 5 and in the previous section, we assume that cells respond deterministically to perturbations that induce a phase response curve. While this approximation may be valid for small perturbations, in reality stochastic noise will ensure that no two cells, even if they start from identical initial conditions, will experience the same phase change from a perturbation. A possible extension to the population-level phase response curve would be incorporating the notion of a phase response *distribution* (PRD), initially

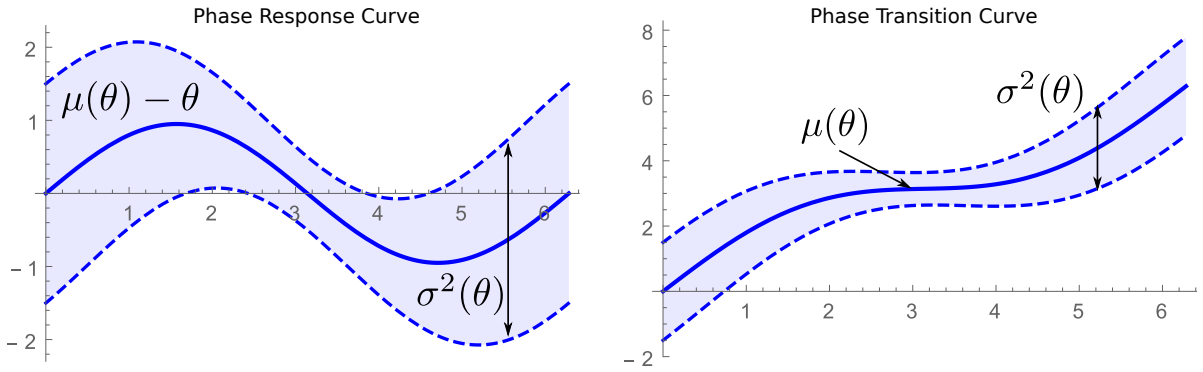


Figure 7.5: Definition of the phase response distribution. Inherent stochasticity in the response of a population of oscillators to stimulus could be captured by allowing a phase-dependent variance to be included in the phase response.

postulated by An *et al.*, 2013 [121]. Single-cell phase response curves have also been tabulated by Pulivarthy *et al.*, 2007, showing variance in the expected phase following perturbation.

A theoretical treatment of the phase response distribution might start with by the phase transition curve as mapping each initial phase to a distribution of perturbed phases $g(\theta_i) = \mathcal{WN}(\mu(\theta_i), \sigma^2(\theta_i); \theta_f)$. A schematic of this is shown in figure 7.5. The phase-transition function g can therefore be imagined as a two-dimensional probability density function, in which $g(x, y)$ represents the probability of an oscillator with initial phase x being transformed to a final phase y. We will call this a phase transition surface, which is useful in calculating the oscillator probabilities following perturbation.

As an example, we take an initial phase population $p(\theta_i) = \mathcal{WN}(\pi/2, 0.5)$ as shown in figure 7.6. In order to find the weighted probability of an oscillator at each initial phase θ_i being mapped to a final phase θ_f, we multiply the initial population $p(\theta_i)$ by the phase transition surface $g(\theta_i, \theta_f)$ to find the weighted phase transition surface, as shown in figure 7.7.

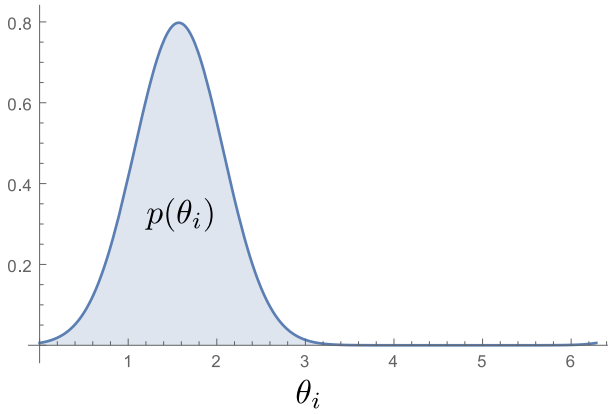


Figure 7.6: Example probability density function for an initial population of oscillators. Taken from a wrapped normal distribution with $\mu = \pi/2$ and $\sigma^2 = 0.5$.

From the weighted phase transition surface, finding the marginal distribution with respect to θ_i is relatively straightforward, and returns the initial phase density prior to perturbation:

$$\begin{aligned}\int_0^{2\pi} p(\theta_i) g(\theta_i, \theta_f)\, d\theta_f &= p(\theta_i) \int_0^{2\pi} \mathcal{WN}(\mu(\theta_i), \sigma^2(\theta_i); \theta_f)\, d\theta_f \\ &= p(\theta_i),\end{aligned}$$

since by definition the area under any wrapped normal distribution from $(0, 2\pi)$ is 1.

Finding the marginal distribution with respect to θ_f, however, is considerably more difficult. It can certainly be obtained numerically, as demonstrated in figure 7.8. The concept of phase response distributions has profound implications for entrainment and synchrony. For instance, in a traditional phase-only model with a phase response curve, oscillators will synchronize at the fixed points defined where the PRC crosses 0 with a negative slope. With a sufficiently small variance in the phase response distribution at a fixed point, relatively good entrainment will likely still be possible. However, for a critical level of variance, it will likely be the case that

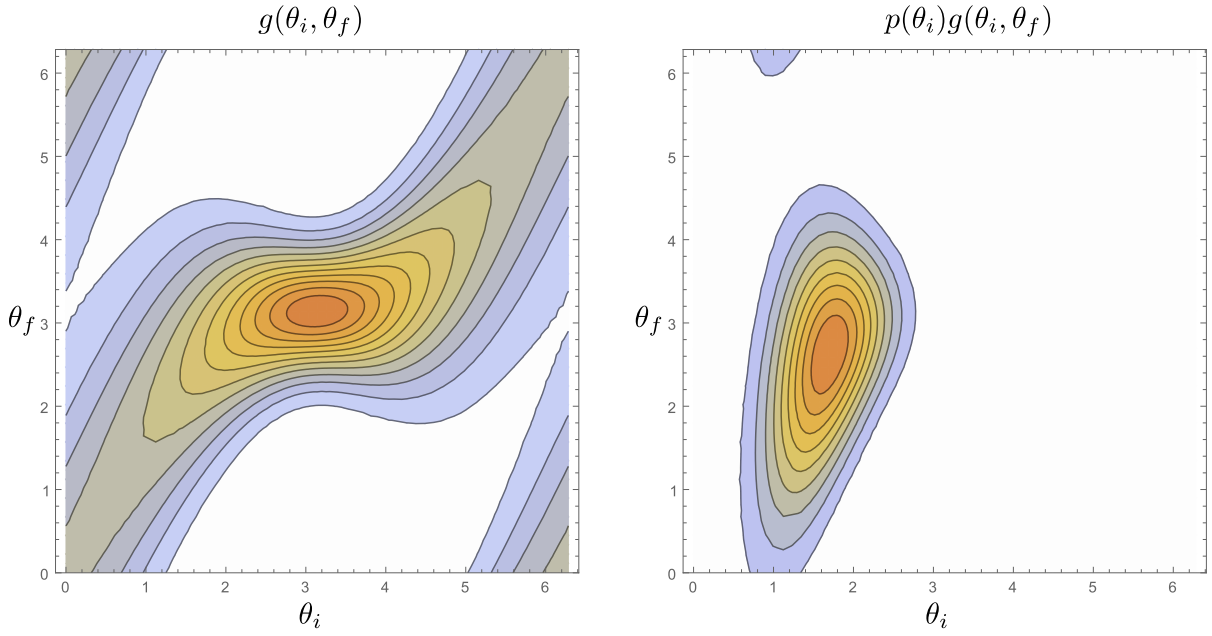


Figure 7.7: Phase transition surfaces. The phase transition surface $g(\theta_i, \theta_f)$ and weighted phase transition surface $p(\theta_i)g(\theta_i, \theta_f)$. From the weighted phase transition surface, integrating in the vertical direction (marginal with respect to θ_i) returns the phase probability density prior to perturbation ($p(\theta_i)$). Integrating in the horizontal direction returns the perturbed phase probability (denoted $\hat{p}(\theta_f)$ in chapter 5).

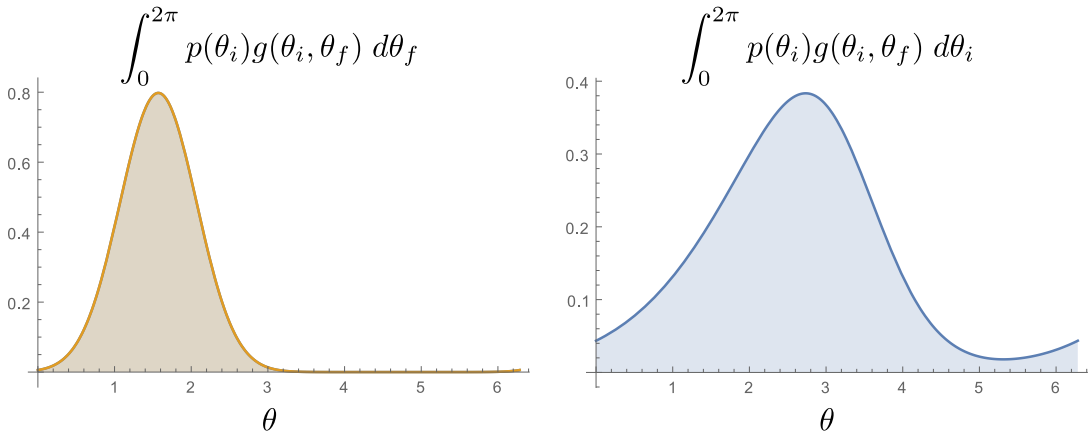


Figure 7.8: Marginal distributions of the weighted phase transition surface. Calculating the marginal distributions of the weighted phase transition surface numerically. The marginal with respect to the final phase (left) returns the phase probability density prior to perturbation. The marginal with respect to the initial phase (right) returns the probability density after perturbation.

entrainment to a given signal is no longer possible.

Finding an analytical solution to

$$\int_0^{2\pi} p(\theta_i)g(\theta_i,\theta_f)\,d\theta_i,$$

might be accomplished by using a series expansion of the mean and variance of the phase response distribution. Such a solution would allow the relationship between the variance and derivative of the PRD at the fixed points to be determined. Even if an analytical solution was not possible, a numerical investigation into the effect of noise in entraining signals would prove useful in finding threshold levels of single-cell noise that would still permit synchronization.

Quantifying phase response distributions

In order for the above analysis to be applicable to experimental systems, we must also develop a suitable method for determining the phase response distribution from single-cell data. This turns out to be a surprisingly difficult proposition, since many forms of error combine to reduce our ability to measure the mean and variance of the phase response. Specifically, localizing a point on the phase response distribution from single-cell data introduces error in both the x-axis (phase before the pulse) as well as the y-axis (phase after the pulse). Since the phase of any given cell is free to wander in a Brownian-type fashion, errors in both directions can be quite high.

One method for measuring the instantaneous phase of a nonstationary signal is the continuous wavelet transform, described in section 7.1.2. Another uses the Hilbert transform, which produces an analytic representation of the initial signal in the complex plane. The phase of the signal can then be obtained from the argument of the Hilbert transform, and used with a weighted linear regression to find an estimate of the phase just before and after perturbation.

To demonstrate such a method on *in silico* data, model 1 was simulated stochastically at a range of different volumes. Perturbations were applied at phases distributed across $(0, 2\pi)$, which consisted of 28% reduction in the vdp parameter (Per *mRNA* degradation) for 6.4 hrs. Control trajectories were simulated in the same fashion, but did not receive the pulse. An example plot of the control and reference trajectories for are shown in figure 7.9, showing the period and amplitude variability due to stochastic noise.

Example phase response distributions are shown for two volumes, $\Omega = 4897$, and $\Omega = 530$, in figure 7.10. Measurement errors seem to dominate the phase response distributions, such that even with 100 data points it would likely be difficult to ascribe

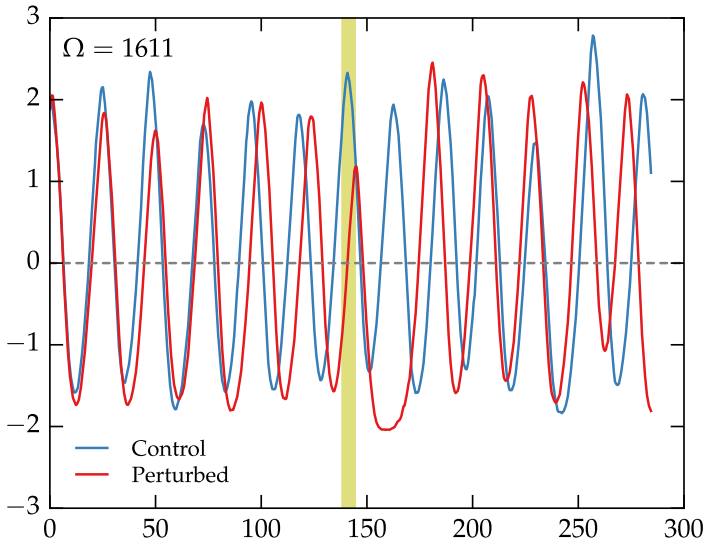
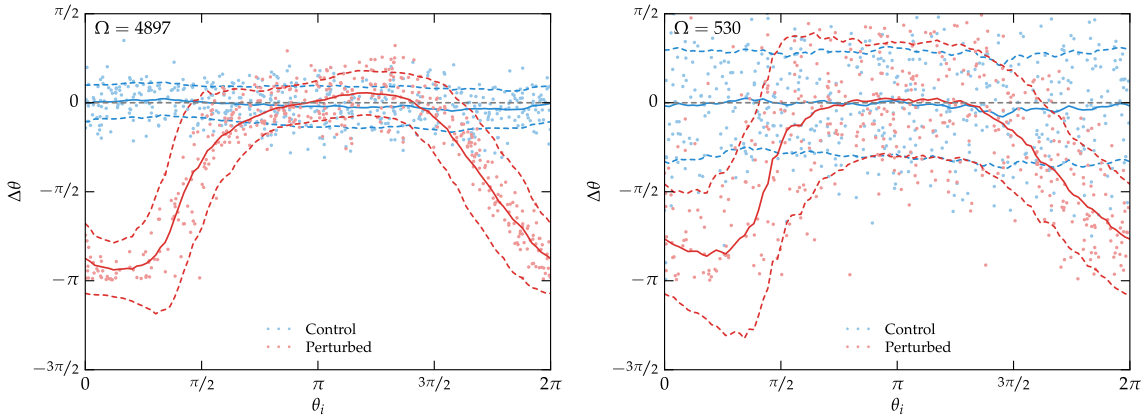


Figure 7.9: Example time-series measurements of noisy oscillators undergoing a phase response. Simulated bioluminescence traces demonstrate the effect of a perturbation that induces a phase change. A control trajectory (blue) does not receive the pulse, but its phase wanders in time due to stochastic fluctuations. The perturbed trajectory (red) undergoes a transient change in amplitude due to the perturbation before settling back to the noisy limit cycle.



(a) $\Omega = 4897$ **(b)** $\Omega = 530$

Figure 7.10: Measured phase response distributions. Phase response distributions as determined for a case with low noise **(a)** and high noise **(b)**. Means and standard deviations (solid and dotted lines, respectively) are calculated from a moving window. In both cases measurement errors, as noticed by error in the control trajectory, will likely make the determination of phase-dependent changes in variance difficult.

statistical significance to phase-dependent variance in the PRD. Regardless, there may be ways of improving this quantification, and estimating the significance of any trends via Bayesian regression. Errors in measuring θ_i and θ_f could be estimated from the combined variance of the control samples.

7.3 Improving identifiability of circadian models by minimizing energy usage

In chapter 3, we describe a means of fitting optimal model parameters repeatedly to bootstrapped experimental data. In using such a method, it became apparent that there were often structural identifiability issues in many circadian models, specifically with regard to protein turnover rate. For example, it was often possible to capture a similar sinusoidal profile both with low degradation and low production

rates, as well as high production and high degradation rates. In techniques such as metabolic flux analysis, a large number of unknowns are optimized by assuming an organism has evolved towards some metabolic goal: for instance optimal growth or minimum energy consumption. Such an approach might also be useful in optimizing circadian models, since actual genetic circuits would likely not use unrealistically high protein turnover rates. By assigning an energy cost to the production of each molecule in the reaction network, the overall "cost" of running the model could be evaluated and minimized alongside the deviation from measured data. A parameter to determine the relative weight of each term would be required, which could be tuned to improve model identifiability while maintaining close fits to experimental data.

7.4 Conclusion

This dissertation has focused on gaining further insight into the dynamics underlying gene regulation through the use of mathematical and computational methods. In one respect, circadian rhythms are an important gene regulatory network with several promising routes by which knowledge might be translated into practical applications. Perhaps more importantly, however, circadian rhythms serve as an excellent example of a genetic circuit for which systems-level dynamics play a critical role in their function. By studying how circadian networks are regulated, we may gain more insight into the biological design principles behind all transcriptional networks. Such insight might be translated into synthetic biology applications, metabolic networks, or other transcriptional systems.

Additionally, the study of circadian rhythms demonstrates the strides that can be taken by the effective collaboration of experimental and computational investigators.

Because the overlap in shared skills is often very little, it can be difficult for computational researchers to judge the quality of experimental work; just as it can be difficult for experimentalists to judge the methodology and reliability of mathematical models. Collaborations are therefore essential not only in generating new data, but also in successfully navigating the voluminous literature on circadian systems. As a result, it is almost assured that the major breakthroughs in understanding and reproducing the incredibly complex nature of gene regulation will come from teams that embrace the interdisciplinary nature of such systems.

Bibliography

[1] A. Relógio, P. O. Westermark, T. Wallach, K. Schellenberg, A. Kramer, and H. Herzel, "Tuning the mammalian circadian clock: robust synergy of two loops," *PLoS Comput. Biol.*, vol. 7, no. 12, p. e1002309, 2011.

[2] J.-C. Leloup and A. Goldbeter, "Toward a detailed computational model for the mammalian circadian clock," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 12, pp. 7051–7056, 2003.

[3] B. Novák and J. J. Tyson, "Design principles of biochemical oscillators," *Nat. Rev. Mol. Cell Biol.*, vol. 9, no. 12, pp. 981–991, 2008.

[4] R. J. Field and R. M. Noyes, "Oscillations in chemical systems. iv. limit cycle behavior in a model of a real chemical reaction," *J. Chem. Phys.*, vol. 60, no. 5, p. 1877, 1974.

[5] J. C. Dunlap, J. J. Loros, and P. J. DeCoursey, *Chronobiology: biological timekeeping*. Sinauer Associates, Inc., 2009.

[6] R. S. Edgar, E. W. Green, Y. Zhao, G. van Ooijen, M. Olmedo, X. Qin, Y. Xu, M. Pan, U. K. Valekunja, K. A. Feeney, E. S. Maywood, M. H. Hastings, N. S. Baliga, M. Merrow, A. J. Millar, C. H. Johnson, C. P. Kyriacou, J. S. O'Neill, and A. B. Reddy, "Peroxiredoxins are conserved markers of circadian rhythms," *Nature*, vol. 485, no. 7399, pp. 459–464, 2012.

[7] M. R. Ralph, R. G. Foster, F. C. Davis, and M. Menaker, "Transplanted suprachiasmatic nucleus determines circadian period," *Science*, vol. 247, no. 4945, pp. 975–978, 1990.

[8] S. J. Aton, C. S. Colwell, A. J. Harmar, J. Waschek, and E. D. Herzog, "Vasoactive intestinal polypeptide mediates circadian rhythmicity and synchrony in mammalian clock neurons," *Nat. Neurosci.*, vol. 8, no. 4, pp. 476–483, 2005.

[9] J. Bass and J. S. Takahashi, "Circadian integration of metabolism and energetics," *Science*, vol. 330, no. 6009, pp. 1349–1354, 2010.

[10] S.-H. Yoo, S. Yamazaki, P. L. Lowrey, K. Shimomura, C. H. Ko, E. D. Buhr, S. M. Siepka, H.-K. Hong, W. J. Oh, O. J. Yoo, M. Menaker, and J. S. Takahashi, "PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 15, pp. 5339–5346, 2004.

[11] S. Yamazaki, R. Numano, M. Abe, A. Hida, R. Takahashi, M. Ueda, G. D. Block, Y. Sakaki, M. Menaker, and H. Tei, "Resetting central and peripheral circadian oscillators in transgenic rats," *Science*, vol. 288, no. 5466, pp. 682–685, 2000.

[12] B. Marcheva, K. M. Ramsey, E. D. Buhr, Y. Kobayashi, H. Su, C. H. Ko, G. Ivanova, C. Omura, S. Mo, M. H. Vitaterna, J. P. Lopez, L. H. Philipson, C. A. Bradfield, S. D. Crosby, L. JeBailey, X. Wang, J. S. Takahashi, and J. Bass, "Disruption of the clock components clock and bmal1 leads to hypoinsulinaemia and diabetes," *Nature*, vol. 466, no. 7306, pp. 627–631, 2010.

[13] S.-Q. Shi, T. S. Ansari, O. P. McGuinness, D. H. Wasserman, and C. H. Johnson, "Circadian disruption leads to insulin resistance and obesity," *Curr. Biol.*, vol. 23, no. 5, pp. 372–381, 2013.

[14] A. Kohsaka, A. D. Laposky, K. M. Ramsey, C. Estrada, C. Joshu, Y. Kobayashi, F. W. Turek, and J. Bass, "High-fat diet disrupts behavioral and molecular circadian rhythms in mice," *Cell Metab.*, vol. 6, no. 5, pp. 414–421, 2007.

[15] M. Hatori, C. Vollmers, A. Zarrinpar, L. DiTacchio, E. A. Bushong, S. Gill, M. Leblanc, A. Chaix, M. Joens, J. A. J. Fitzpatrick, M. H. Ellisman, and S. Panda, "Time-restricted feeding without reducing caloric intake prevents metabolic diseases in mice fed a high-fat diet," *Cell Metab.*, vol. 15, no. 6, pp. 848–860, 2012.

[16] C. H. Ko and J. S. Takahashi, "Molecular components of the mammalian circadian clock," *Hum. Mol. Genet.*, vol. 15, no. S2, pp. R271–R277, 2006.

[17] C. Lee, J.-P. Etchegaray, F. R. Cagampang, A. S. Loudon, and S. M. Reppert, "Posttranslational mechanisms regulate the mammalian circadian clock," *Cell*, vol. 107, no. 7, pp. 855–867, 2001.

[18] B. Ananthasubramaniam and H. Herzel, "Positive feedback promotes oscillations in negative feedback loops," *PLoS One*, vol. 9, no. 8, p. e104761, 2014.

[19] S. Panda, J. B. Hogenesch, and S. A. Kay, "Circadian rhythms from flies to human," *Nature*, vol. 417, no. May, pp. 329–335, 2002.

[20] M. H. Vitaterna, D. P. King, A. M. Chang, J. M. Kornhauser, P. L. Lowrey, J. D. McDonald, W. F. Dove, L. H. Pinto, F. W. Turek, and J. S. Takahashi, "Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior," *Science*, vol. 264, no. 5159, pp. 719–725, 1994.

[21] G. T. van der Horst, M. Muijtjens, K. Kobayashi, R. Takano, S. Kanno, M. Takao, J. de Wit, A. Verkerk, A. P. Eker, D. van Leenen, R. Buijs, D. Bootsma, J. H. Hoeijmakers, and A. Yasui, "Mammalian cry1 and cry2 are essential for maintenance of circadian rhythms," *Nature*, vol. 398, no. 6728, pp. 627–630, 1999.

[22] E. E. Zhang, A. C. Liu, T. Hirota, L. J. Miraglia, G. Welch, P. Y. Pongsawakul, X. Liu, A. Atwood, J. W. Huss, J. Janes, A. I. Su, J. B. Hogenesch, and S. A. Kay, "A genome-wide RNAi screen for modifiers of the circadian clock in human cells," *Cell*, vol. 139, no. 1, pp. 199–210, 2009.

[23] T. Hirota, W. G. Lewis, A. C. Liu, J. W. Lee, P. G. Schultz, and S. A. Kay, "A chemical biology approach reveals period shortening of the mammalian circadian clock by specific inhibition of GSK-3beta," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 52, pp. 20746–20751, 2008.

[24] T. Noguchi, L. L. Wang, and D. K. Welsh, "Fibroblast PER2 circadian rhythmicity depends on cell density," *J. Biol. Rhythms*, vol. 28, no. 3, pp. 183–192, 2013.

[25] D. K. Welsh, S.-H. Yoo, A. C. Liu, J. S. Takahashi, and S. A. Kay, "Bioluminescence imaging of individual fibroblasts reveals persistent, independently phased circadian rhythms of clock gene expression," *Curr. Biol.*, vol. 14, no. 24, pp. 2289–2295, 2004.

[26] M. Izumo, T. R. Sato, M. Straume, and C. H. Johnson, "Quantitative analyses of circadian gene expression in mammalian cell cultures," *PLoS Comput. Biol.*, vol. 2, no. 10, p. e136, 2006.

[27] A. T. Winfree, *The Geometry of Biological Time*, vol. 12 of *Interdisciplinary Applied Mathematics*. New York, New York: Springer, 2001.

[28] R. P. Dickinson and R. J. Gelinas, "Sensitivity analysis of ordinary differential equation systems – a direct method," *J. Comput. Phys.*, vol. 21, no. 2, pp. 123–143, 1976.

[29] W. F. Feehery, J. E. Tolsma, and P. I. Barton, "Efficient sensitivity analysis of large-scale differential-algebraic systems," *Appl. Numer. Math.*, vol. 25, no. 1, pp. 41–54, 1997.

[30] M. A. Kramer, H. Rabitz, and J. M. Calo, "Sensitivity analysis of oscillatory systems," *Appl. Math. Model.*, vol. 8, no. 5, pp. 328–340, 1984.

[31] A. K. Wilkins, B. Tidor, J. White, and P. I. Barton, "Sensitivity analysis for oscillating dynamical systems," *SIAM J. Sci. Comput.*, vol. 31, no. 4, p. 2706, 2009.

[32] S. R. Taylor, R. Gunawan, L. R. Petzold, and F. J. Doyle III, "Sensitivity measures for oscillating systems: application to mammalian circadian gene network," *IEEE Trans. Automat. Contr.*, vol. 53, no. January, pp. 177–188, 2008.

[33] E. O. Voit, H. A. Martens, and S. W. Omholt, "150 years of the mass action law," *PLoS Comput. Biol.*, vol. 11, no. 1, pp. 1–7, 2015.

[34] K. A. Johnson and R. S. Goody, "The original Michaelis constant: translation of the 1913 Michaelis-Menten paper," *Biochemistry*, vol. 50, no. 39, pp. 8264–8269, 2011.

[35] A. Tzafriri and E. Edelman, "The total quasi-steady-state approximation is valid for reversible enzyme kinetics," *J. Theor. Biol.*, vol. 226, no. 3, pp. 303–313, 2004.

[36] A. V. Hill, "The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves," *J. Physiol.*, vol. 40, pp. i—-vii, 1910.

[37] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, "Global quantification of mammalian gene expression control," *Nature*, vol. 473, pp. 337–342, 2011.

[38] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Phys. Chem.*, vol. 84, no. 25, pp. 2340–2361, 1977.

[39] K. R. Sanft, D. T. Gillespie, and L. R. Petzold, "Legitimacy of the stochastic michaelis-menten approximation," *IET Syst. Biol.*, vol. 5, no. 1, p. 58, 2011.

[40] D. Whitley, "An overview of evolutionary algorithms: practical issues and common pitfalls," *Inf. Softw. Technol.*, vol. 43, no. 14, pp. 817–831, 2001.

[41] B. C. Goodwin, "Oscillatory behavior in enzymatic control processes," *Adv. Enzyme Regul.*, vol. 3, pp. 425–438, 1965.

[42] A. Goldbeter, "A model for circadian oscillations in the drosophila period protein (PER)," *Proc. Biol. Sci.*, vol. 261, no. 1362, pp. 319–324, 1995.

[43] D. B. Forger and C. S. Peskin, "A detailed predictive model of the mammalian circadian clock," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 25, pp. 14806–14811, 2003.

[44] H. P. Mirsky, A. C. Liu, D. K. Welsh, S. A. Kay, and F. J. Doyle III, "A model of the cell-autonomous mammalian circadian clock," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 27, pp. 11107–11112, 2009.

[45] J. K. Kim and D. B. Forger, "A mechanism for robust circadian timekeeping via stoichiometric balance," *Mol. Syst. Biol.*, vol. 8, no. 630, pp. 1–14, 2012.

[46] D. B. Forger and C. S. Peskin, "Stochastic simulation of the mammalian circadian clock," *PNAS*, vol. 102, no. 2, pp. 321–324, 2005.

[47] T.-L. To, M. A. Henson, E. D. Herzog, and F. J. Doyle III, "A molecular model for intercellular synchronization in the mammalian circadian clock," *Biophys. J.*, vol. 92, no. 11, pp. 3792–3803, 2007.

[48] Z. Chen, S.-H. Yoo, and J. S. Takahashi, "Small molecule modifiers of circadian clocks," *Cell. Mol. Life Sci.*, vol. 70, no. 16, pp. 2985–2998, 2013.

[49] T. Hirota, J. W. Lee, P. C. St. John, M. Sawa, K. Iwaisako, T. Noguchi, P. Y. Pongsawakul, T. Sonntag, D. K. Welsh, D. A. Brenner, F. J. Doyle III, P. G. Schultz, and S. A. Kay, "Identification of small molecule activators of cryptochrome," *Science*, vol. 337, no. 6098, pp. 1094–1097, 2012.

[50] E. V. McCarthy, J. E. Baggs, J. M. Geskes, J. B. Hogenesch, and C. B. Green, "Generation of a novel allelic series of cryptochrome mutants via mutagenesis reveals residues involved in protein-protein interaction and CRY2-specific repression," *Mol. Cell. Biol.*, vol. 29, no. 20, pp. 5465–5476, 2009.

[51] K. Miyazaki, M. Mesaki, and N. Ishida, "Nuclear entry mechanism of rat PER2 (rPER2): role of rPER2 in nuclear localization of CRY protein," *Mol. Cell. Biol.*, vol. 21, no. 19, pp. 6651–6659, 2001.

[52] K. Kume, M. J. Zylka, S. Sriram, L. P. Shearman, D. R. Weaver, X. Jin, E. S. Maywood, M. H. Hastings, and S. M. Reppert, "mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop," *Cell*, vol. 98, no. 2, pp. 193–205, 1999.

[53] Y. Lee, R. Chen, H.-m. Lee, and C. Lee, "Stoichiometric relationship among clock proteins determines robustness of circadian rhythms," *J. Biol. Chem.*, vol. 286, no. 9, pp. 7033–7042, 2011.

[54] N. A. Cookson, L. S. Tsimring, and J. Hasty, "The pedestrian watchmaker: genetic clocks from engineered oscillators," *FEBS Lett.*, vol. 583, no. 24, pp. 3931–3937, 2009.

[55] N. Kurabayashi, T. Hirota, M. Sakai, K. Sanada, and Y. Fukada, "DYRK1A and glycogen synthase kinase 3beta, a dual-kinase mechanism directing proteasomal degradation of CRY2 for circadian timekeeping," *Mol. Cell. Biol.*, vol. 30, no. 7, pp. 1757–1768, 2010.

[56] E. Vielhaber, E. Eide, A. Rivers, Z.-H. Gao, and D. M. Virshup, "Nuclear entry of the circadian regulator mPER1 is controlled by mammalian casein kinase Ie," *Mol. Cell. Biol.*, vol. 20, no. 13, pp. 4888–4899, 2000.

[57] E. A. Griffin Jr., D. Staknis, and C. J. Weitz, "Light-independent role of CRY1 and CRY2 in the mammalian circadian clock," *Science*, vol. 286, no. 5440, pp. 768–771, 1999.

[58] R. Yang, M. Rodriguez-Fernandez, P. C. St. John, and F. J. Doyle III, "8 –systems biology," in *Model. Methodol. Physiol. Med. (Second Ed.)*, ch. 8, pp. 159–187, Elsevier, 2014.

[59] J. Andersson, B. Houska, and M. Diehl, "Towards a computer algebra system with automatic differentiation for use with object-oriented modelling languages," *3rd Int. Work. Equation-Based Object-Oriented Model. Lang. Tools*, pp. 99–105, 2010.

[60] A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. A. N. E. Shumaker, and C. S. Woodward, "Sundials : suite of nonlinear and differential / algebraic equation solvers," *ACM Trans. Math. Softw.*, vol. 31, no. 3, pp. 363–396, 2005.

[61] A. C. Liu, D. K. Welsh, C. H. Ko, H. G. Tran, E. E. Zhang, A. A. Priest, E. D. Buhr, O. Singer, K. Meeker, I. M. Verma, F. J. Doyle III, J. S. Takahashi, and S. A. Kay, "Intercellular coupling confers robustness against mutations in the SCN circadian clock network," *Cell*, vol. 129, no. 3, pp. 605–616, 2007.

[62] S. Nangle, W. Xing, and N. Zheng, "Crystal structure of mammalian cryptochrome in complex with a small molecule competitor of its ubiquitin ligase," *Cell Res.*, pp. 1–3, 2013.

[63] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna, "Universally sloppy parameter sensitivities in systems biology models," *PLoS Comput. Biol.*, vol. 3, no. 10, pp. 1871–1878, 2007.

[64] D. B. Kell and J. D. Knowles, "The role of modeling in systems biology," in *Syst. Model. Cell. Biol. from concepts to nuts bolts*, pp. 3–18, MIT Press, 2006.

[65] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer, "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood," *Bioinformatics*, vol. 25, no. 15, pp. 1923–1929, 2009.

[66] M. Nihtilä and J. Virkkunen, "Practical identifiability of growth and substrate consumption models," *Biotechnol. Bioeng.*, vol. 19, no. 12, pp. 1831–1850, 1977.

[67] J. E. Jiménez-Hornero, I. M. Santos-Dueñas, and I. García-García, "Optimization of biotechnological processes. the acetic acid fermentation. part ii: practical identifiability analysis and parameter estimation," *Biochem. Eng. J.*, vol. 45, no. 1, pp. 7–21, 2009.

[68] A. Holmberg, "On the practical identifiability of microbial growth models incorporating Michaelis-Menten type nonlinearities," *Math. Biosci.*, vol. 62, no. 1, pp. 23–43, 1982.

[69] M. Joshi, A. Seidel-Morgenstern, and A. Kremling, "Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems," *Metab. Eng.*, vol. 8, no. 5, pp. 447–455, 2006.

[70] M. E. Hughes, L. DiTacchio, K. R. Hayes, C. Vollmers, S. Pulivarthy, J. E. Baggs, S. Panda, and J. B. Hogenesch, "Harmonics of circadian gene transcription in mammals," *PLoS Genet.*, vol. 5, no. 4, p. e1000442, 2009.

[71] T. Hirota, J. W. Lee, W. G. Lewis, E. E. Zhang, G. Breton, X. Liu, M. Garcia, E. C. Peters, J.-P. Etchegaray, D. Traver, P. G. Schultz, and S. A. Kay, "High-throughput chemical screen identifies a novel potent modulator of cellular circadian rhythms and reveals CKIα as a clock regulatory kinase," *PLoS Biol.*, vol. 8, no. 12, p. e1000559, 2010.

[72] L. T. Biegler, *Nonlinear Programming: concepts, Algorithms, and Applications to Chemical Processes*. SIAM, 2010.

[73] H. G. Bock, E. Kostina, and J. P. Schloder, "Numerical methods for parameter estimation in nonlinear differential algebraic equations," *GAMM Mitteilungen*, vol. 408, no. 2, pp. 376 – 408, 2007.

[74] C. A. Floudas, *Nonlinear and Mixed-Integer Optimization*. New York, New York: Oxford University Press, 1995.

[75] A. Wachter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Math. Program.*, vol. 106, no. 1, pp. 25–57, 2006.

[76] HSL, "A collection of fortran codes for large scale scientific computation." http://www.hsl.rl.ac.uk/, 2011.

[77] J. Andersson, *A General-Purpose Software Framework for Dynamic Optimization*. PhD thesis, Arenberg Doctoral School, KU Leuven, 2013.

[78] E. Jones, T. Oliphant, and P. Peterson, "Scipy: open source scientific tools for python." http://www.scipy.org, 2001.

[79] E. G. Bure and E. Rosenwasser, "The study of the sensitivity of oscillatory systems," *Autom. Remote Control*, vol. 7, pp. 1045–1052, 1974.

[80] D. E. Zak, G. E. Gonye, J. S. Schwaber, and F. J. Doyle III, "Importance of input perturbations and stochastic gene expression in the reverse engineering

of genetic regulatory networks: insights from an identifiability analysis of an in silico network," *Genome Res.*, vol. 13, no. 11, pp. 2396–2405, 2003.

[81] P. C. St John and F. J. Doyle III, "Estimating confidence intervals in predicted responses for oscillatory biological models," *BMC Syst. Biol.*, vol. 7, no. 1, p. 71, 2013.

[82] J. Bass, "Circadian topology of metabolism," *Nature*, vol. 491, no. 7424, pp. 348–356, 2012.

[83] K. Griffett and T. P. Burris, "The mammalian clock and chronopharmacology," *Bioorg. Med. Chem. Lett.*, vol. 23, no. 7, pp. 1929–1934, 2013.

[84] R. Ye, C. P. Selby, N. Ozturk, Y. Annayev, and A. Sancar, "Biochemical analysis of the canonical model for the mammalian circadian clock," *J. Biol. Chem.*, vol. 286, no. 29, pp. 25891–25902, 2011.

[85] K. Yagita, F. Tamanini, M. Yasuda, J. H. J. Hoeijmakers, G. T. J. van der Horst, and H. Okamura, "Nucleocytoplasmic shuttling and mCRY-dependent inhibition of ubiquitylation of the mPER2 clock protein," *EMBO J.*, vol. 21, no. 6, pp. 1301–1314, 2002.

[86] S. Reischl, K. Vanselow, P. O. Westermark, N. Thierfelder, B. Maier, H. Herzel, and A. Kramer, "Beta-TrCP1-mediated degradation of PERIOD2 is essential for circadian dynamics," *J. Biol. Rhythms*, vol. 22, no. 5, pp. 375–386, 2007.

[87] A. Takano, Y. Isojima, and K. Nagai, "Identification of mPer1 phosphorylation sites responsible for the nuclear entry," *J. Biol. Chem.*, vol. 279, no. 31, pp. 32578–32585, 2004.

[88] L. Busino, F. Bassermann, A. Maiolica, C. Lee, P. M. Nolan, S. I. H. Godinho, G. F. Draetta, and M. Pagano, "SCFFbxl3 controls the oscillation of the circadian clock by directing the degradation of cryptochrome proteins," *Science*, vol. 316, no. 5826, pp. 900–904, 2007.

[89] S. I. H. Godinho, E. S. Maywood, L. Shaw, V. Tucci, A. R. Barnard, L. Busino, M. Pagano, R. Kendall, M. M. Quwailid, M. R. Romero, J. O'neill, J. E. Chesham, D. Brooker, Z. Lalanne, M. H. Hastings, and P. M. Nolan, "The after-hours mutant reveals a role for Fbxl3 in determining mammalian circadian period," *Science*, vol. 316, no. 5826, pp. 897–900, 2007.

[90] S. M. Siepka, S.-H. Yoo, J. Park, W. Song, V. Kumar, Y. Hu, C. Lee, and J. S. Takahashi, "Circadian mutant overtime reveals F-box protein FBXL3 regulation of cryptochrome and period gene expression," *Cell*, vol. 129, no. 5, pp. 1011–1023, 2007.

[91] Y. Lee and E.-K. Kim, "AMP-activated protein kinase as a key molecular link between metabolism and clockwork," *Exp. Mol. Med.*, vol. 45, no. 7, p. e33, 2013.

[92] M. Gallego, E. J. Eide, M. F. Woolf, D. M. Virshup, and D. B. Forger, "An opposite role for tau in circadian rhythms revealed by mathematical modeling," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 28, pp. 10618–10623, 2006.

[93] E. S. Maywood, J. E. Chesham, Q.-J. Meng, P. M. Nolan, A. S. I. Loudon, and M. H. Hastings, "Tuning the period of the mammalian circadian clock: additive and independent effects of CK1ϵTau and Fbxl3Afh mutations on mouse circadian behavior and molecular pacemaking," *J. Neurosci.*, vol. 31, no. 4, pp. 1539–1544, 2011.

[94] Q.-J. Meng, L. Logunova, E. S. Maywood, M. Gallego, J. Lebiecki, T. M. Brown, M. Sládek, A. S. Semikhodskii, N. R. J. Glossop, H. D. Piggins, J. E. Chesham, D. A. Bechtold, S.-h. Yoo, J. S. Takahashi, D. M. Virshup, R. P. Boot-Handford, M. H. Hastings, and A. S. I. Loudon, "Setting clock speed in mammals: the CK1 epsilon tau mutation in mice accelerates circadian pacemakers by selectively destabilizing PERIOD proteins," *Neuron*, vol. 58, no. 1, pp. 78–88, 2008.

[95] K. Vanselow, J. T. Vanselow, P. O. Westermark, S. Reischl, B. Maier, T. Korte, A. Herrmann, H. Herzel, A. Schlosser, and A. Kramer, "Differential effects of PER2 phosphorylation: molecular basis for the human familial advanced sleep phase syndrome (FASPS)," *Genes Dev.*, vol. 20, no. 19, pp. 2660–2672, 2006.

[96] P. C. St. John, T. Hirota, S. A. Kay, and F. J. Doyle III, "Spatiotemporal separation of PER and CRY posttranslational regulation in the mammalian circadian clock," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 5, pp. 2040–2045, 2014.

[97] R. Gunawan and F. J. Doyle III, "Isochron-based phase response analysis of circadian rhythms," *Biophys. J.*, vol. 91, no. 6, pp. 2131–2141, 2006.

[98] K. Ohsaki, K. Oishi, Y. Kozono, K. Nakayama, K. I. Nakayama, and N. Ishida, "The role of beta-TrCP1 and beta-TrCP2 in circadian rhythm generation by mediating degradation of clock protein PER2," *J. Biochem.*, vol. 144, no. 5, pp. 609–618, 2008.

[99] H. Lee, R. Chen, Y. Lee, S. Yoo, and C. Lee, "Essential roles of CKIdelta and CKIepsilon in the mammalian circadian clock," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 50, pp. 21359–21364, 2009.

[100] A. Hirano, K. Yumimoto, R. Tsunematsu, M. Matsumoto, M. Oyama, H. Kozuka-Hata, T. Nakagawa, D. Lanjakornsiripan, K. I. Nakayama, and Y. Fukada, "FBXL21 regulates oscillation of the circadian clock through ubiquitination and stabilization of cryptochromes," *Cell*, vol. 152, no. 5, pp. 1106–1118, 2013.

[101] S.-H. Yoo, J. A. Mohawk, S. M. Siepka, Y. Shan, S. K. Huh, H.-K. Hong, I. Kornblum, V. Kumar, N. Koike, M. Xu, J. Nussbaum, X. Liu, Z. Chen, Z. J. Chen, C. B. Green, and J. S. Takahashi, "Competing E3 ubiquitin ligases govern circadian periodicity by degradation of CRY in nucleus and cytoplasm," *Cell*, vol. 152, no. 5, pp. 1091–1105, 2013.

[102] K. A. Lamia, K.-F. Storch, and C. J. Weitz, "Physiological significance of a peripheral tissue circadian clock," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 39, pp. 15172–15177, 2008.

[103] B. Kornmann, O. Schaad, H. Bujard, J. S. Takahashi, and U. Schibler, "System-driven and oscillator-dependent circadian transcription in mice with a conditionally active liver clock," *PLoS Biol.*, vol. 5, no. 2, p. e34, 2007.

[104] C. S. Pittendrigh, W. T. Kyner, and T. Takamura, "The amplitude of circadian oscillations: temperature dependence, latitudinal clines, and the photoperiodic time measurement," *J. Biol. Rhythms*, vol. 6, no. 4, pp. 299–313, 1991.

[105] U. Abraham, A. E. Granada, P. O. Westermark, M. Heine, A. Kramer, and H. Herzel, "Coupling governs entrainment range of circadian clocks," *Mol. Syst. Biol.*, vol. 6, no. 438, p. 438, 2010.

[106] H.-C. Chang and L. Guarente, "SIRT1 mediates central circadian control in the SCN by a mechanism that decays with aging," *Cell*, vol. 153, no. 7, pp. 1448–1460, 2013.

[107] C. Ramanathan, H. Xu, S. K. Khan, Y. Shen, P. J. Gitis, D. K. Welsh, J. B. Hogenesch, and A. C. Liu, "Cell type-specific functions of period genes revealed by novel adipocyte and hepatocyte circadian clock models," *PLoS Genet.*, vol. 10, no. 4, p. e1004244, 2014.

[108] J. E. Baggs, T. S. Price, L. DiTacchio, S. Panda, G. A. Fitzgerald, and J. B. Hogenesch, "Network features of the mammalian circadian clock," *PLoS Biol.*, vol. 7, no. 3, p. e52, 2009.

[109] S. R. Pulivarthy, N. Tanaka, D. K. Welsh, L. De Haro, I. M. Verma, and S. Panda, "Reciprocity between phase shifts and amplitude changes in the mammalian circadian clock," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 51, pp. 20356–20361, 2007.

[110] H. Ukai, T. J. Kobayashi, M. Nagano, K.-h. Masumoto, M. Sujino, T. Kondo, K. Yagita, Y. Shigeyoshi, and H. R. Ueda, "Melanopsin-dependent photoperturbation reveals desynchronization underlying the singularity of mammalian circadian clocks," *Nat. Cell Biol.*, vol. 9, no. 11, pp. 1327–1334, 2007.

[111] D. A. Rand, B. V. Shulgin, D. Salazar, and A. J. Millar, "Design principles underlying circadian clocks," *J. R. Soc. Interface*, vol. 1, no. 1, pp. 119–130, 2004.

[112] S. Daan and C. S. Pittendrigh, "A functional analysis of circadian pacemakers in nocturnal rodents," *J. Comp. Physiol. A*, vol. 106, no. 3, pp. 253–266, 1976.

[113] S. R. Taylor, F. J. Doyle III, and L. R. Petzold, "Oscillator model reduction preserving the phase response: application to the circadian clock," *Biophys. J.*, vol. 95, no. 4, pp. 1658–1673, 2008.

[114] B. Pfeuty, Q. Thommen, and M. Lefranc, "Robust entrainment of circadian oscillators requires specific phase response curves," *Biophys. J.*, vol. 100, no. 11, pp. 2557–2565, 2011.

[115] J. K. Kim, Z. P. Kilpatrick, M. R. Bennett, and K. Josić, "Molecular mechanisms that regulate the coupled period of the mammalian circadian clock," *Biophys. J.*, vol. 106, no. 9, pp. 2071–2081, 2014.

[116] M. E. Jewett and R. E. Kronauer, "Refinement of a limit cycle oscillator model of the effects of light on the human circadian pacemaker," *J. Theor. Biol.*, vol. 192, no. 4, pp. 455–465, 1998.

[117] P. Achermann and H. Kunz, "Modeling circadian rhythm generation in the suprachiasmatic nucleus with locally coupled self-sustained oscillators: phase shifts and phase response curves," *J. Biol. Rhythms*, vol. 14, no. 6, pp. 460–468, 1999.

[118] D. R. van der Veen, J. Shao, Y. Xi, L. Li, and G. E. Duffield, "Cardiac atrial circadian rhythms in PERIOD2::LUCIFERASE and per1:luc mice: amplitude and phase responses to glucocorticoid signaling and medium treatment," *PLoS One*, vol. 7, no. 10, p. e47692, 2012.

[119] O. Castejón, A. Guillamon, and G. Huguet, "Phase-amplitude response functions for transient-state stimuli," *J. Math. Neurosci.*, vol. 3, no. 1, p. 13, 2013.

[120] F. van Oosterhout, E. A. Lucassen, T. Houben, H. T. VanderLeest, M. C. Antle, and J. H. Meijer, "Amplitude of the SCN clock enhanced by the behavioral activity rhythm," *PLoS One*, vol. 7, no. 6, p. e39693, 2012.

[121] S. An, R. Harang, K. Meeker, D. Granados-Fuentes, C. A. Tsai, C. Mazuski, J. Kim, F. J. Doyle III, L. R. Petzold, and E. D. Herzog, "A neuropeptide speeds circadian entrainment by reducing intercellular synchrony," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, no. 46, pp. E4355–E4361, 2013.

[122] P. C. St. John, S. R. Taylor, J. H. Abel, and F. J. Doyle III, "Amplitude metrics for cellular circadian bioluminescence reporters," *Biophys. J.*, vol. 107, no. 11, pp. 2712–2722, 2014.

[123] J. Rougemont and F. Naef, "Collective synchronization in populations of globally coupled phase oscillators with drifting frequencies," *Phys. Rev. E*, vol. 73, no. 1, p. 011104, 2006.

[124] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*. Berlin: Springer-Verlag, 1984.

[125] J.-n. Teramae and D. Tanaka, "Robustness of the noise-induced phase synchronization in a general class of limit cycle oscillators," *Phys. Rev. Lett.*, vol. 93, no. 20, p. 204103, 2004.

[126] R. B. Stein, "A theoretical analysis of neuronal variability," *Biophys. J.*, vol. 5, no. 5, pp. 173–194, 1965.

[127] J. Rougemont and F. Naef, "Dynamical signatures of cellular fluctuations and oscillator stability in peripheral circadian clocks," *Mol. Syst. Biol.*, vol. 3, no. 93, p. 93, 2007.

[128] G. S. Chirikjian, "Gaussian distributions and the heat equation," in *Stoch. Model. Inf. Theory, Lie Groups, Vol. 1*, vol. 1, ch. 2, pp. 31–61, Boston: Birkhäuser, 2009.

[129] K. V. Mardia and P. E. Jupp, *Directional Statistics*, vol. 494. John Wiley & Sons, 2009.

[130] K. R. Sanft, S. Wu, M. Roh, J. Fu, R. K. Lim, and L. R. Petzold, "Stochkit2: software for discrete stochastic simulation of biochemical systems with events," *Bioinformatics*, vol. 27, no. 17, pp. 2457–2458, 2011.

[131] T. L. Leise and M. E. Harrington, "Wavelet-based time series analysis of circadian rhythms," *J. Biol. Rhythms*, vol. 26, no. 5, pp. 454–463, 2011.

[132] S. A. Tischkau, J. W. Mitchell, S.-H. Tyan, G. F. Buchanan, and M. U. Gillette, "Ca2+/cAMP response element-binding protein (CREB)-dependent activation of Per1 is required for light-induced signaling in the suprachiasmatic nucleus circadian clock," *J. Biol. Chem.*, vol. 278, no. 2, pp. 718–723, 2003.

[133] J. H. T. Rohling, H. T. VanderLeest, S. Michel, M. J. Vansteensel, and J. H. Meijer, "Phase resetting of the mammalian circadian clock relies on a rapid shift of a small population of pacemaker neurons," *PLoS One*, vol. 6, no. 9, p. e25437, 2011.

[134] A. Balsalobre, F. Damiola, and U. Schibler, "A serum shock induces circadian gene expression in mammalian tissue culture cells," *Cell*, vol. 93, no. 6, pp. 929–937, 1998.

[135] C. J. Guenthner, M. E. Luitje, L. A. Pyle, P. C. Molyneux, J. K. Yu, A. S. Li, T. L. Leise, and M. E. Harrington, "Circadian rhythms of PER2::LUC in individual primary mouse hepatocytes and cultures," *PLoS One*, vol. 9, no. 2, p. e87573, 2014.

[136] E. Nagoshi, C. Saini, C. Bauer, T. Laroche, F. Naef, and U. Schibler, "Circadian gene expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass time to daughter cells," *Cell*, vol. 119, no. 5, pp. 693–705, 2004.

[137] P. O. Westermark, D. K. Welsh, H. Okamura, and H. Herzel, "Quantification of circadian rhythms in single cells," *PLoS Comput. Biol.*, vol. 5, no. 11, p. e1000580, 2009.

[138] T. L. Leise, C. W. Wang, P. J. Gitis, and D. K. Welsh, "Persistent cell-autonomous circadian oscillations in fibroblasts revealed by six-week single-cell imaging of PER2::LUC bioluminescence," *PLoS One*, vol. 7, no. 3, pp. 1–10, 2012.

[139] M. Ravn and H. Uhlig, "On adjusting the hodrick-prescott filter for the frequency of observations," *Rev. Econ. Stat.*, vol. 84, no. 2, pp. 371–376, 2002.

[140] Y. Hua and T. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust.*, vol. 38, no. 5, pp. 814–824, 1990.

[141] T. P. Zieliński and K. Duda, "Frequency and damping estimation methods – an overview," *Metrol. Meas. Syst.*, vol. 18, no. 4, pp. 505–528, 2011.

[142] A. Birmingham, L. M. Selfors, T. Forster, D. Wrobel, C. J. Kennedy, E. Shanks, J. Santoyo-Lopez, D. J. Dunican, A. Long, D. Kelleher, Q. Smith, R. L. Beijersbergen, P. Ghazal, and C. E. Shamu, "Statistical methods for analysis of high-throughput RNA interference screens," *Nat. Methods*, vol. 6, no. 8, pp. 569–575, 2009.

[143] M. Izumo, C. H. Johnson, and S. Yamazaki, "Circadian gene expression in mammalian fibroblasts revealed by real-time luminescence reporting: temperature compensation and damping," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, pp. 16089–16094, 2003.

[144] C. Vollmers, S. Gill, L. DiTacchio, S. R. Pulivarthy, H. D. Le, and S. Panda, "Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 50, pp. 21453–21458, 2009.

[145] D. Yamajuku, T. Inagaki, T. Haruma, S. Okubo, Y. Kataoka, S. Kobayashi, K. Ikegami, T. Laurent, T. Kojima, K. Noutomi, S. Hashimoto, and H. Oda, "Real-time monitoring in three-dimensional hepatocytes reveals that insulin acts as a synchronizer for liver clock," *Sci. Rep.*, vol. 2, p. 439, 2012.

[146] M. Naïmi, C. Arous, and E. Van Obberghen, "Energetic cell sensors: a key to metabolic homeostasis," *Trends Endocrinol. Metab.*, vol. 21, no. 2, pp. 75–82, 2010.

[147] C. Luni and F. J. Doyle III, "Robust multi-drug therapy design and application to insulin resistance in type 2 diabetes," *Int. J. Robust Nonlinear Control*, vol. 21, no. 15, pp. 1730–1741, 2011.

[148] A. Hahn-Windgassen, V. Nogueira, C.-C. Chen, J. E. Skeen, N. Sonenberg, and N. Hay, "Akt activates the mammalian target of rapamycin by regulating cellular ATP level and AMPK activity," *J. Biol. Chem.*, vol. 280, no. 37, pp. 32081–32089, 2005.

[149] R. Cao and K. Obrietan, "mTOR signaling and entrainment of the mammalian circadian clock," *Mol. Cell. Pharmacol.*, vol. 2, no. 4, pp. 125–130, 2010.

[150] K. A. Lamia, U. M. Sachdeva, L. DiTacchio, E. C. Williams, J. G. Alvarez, D. F. Egan, D. S. Vasquez, H. Juguilon, S. Panda, R. J. Shaw, C. B. Thompson, and R. M. Evans, "AMPK regulates the circadian clock by cryptochrome phosphorylation and degradation," *Science*, vol. 326, no. 5951, pp. 437–440, 2009.

[151] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bull. Am. Meteorol. Soc.*, vol. 79, no. 1, pp. 61–78, 1998.